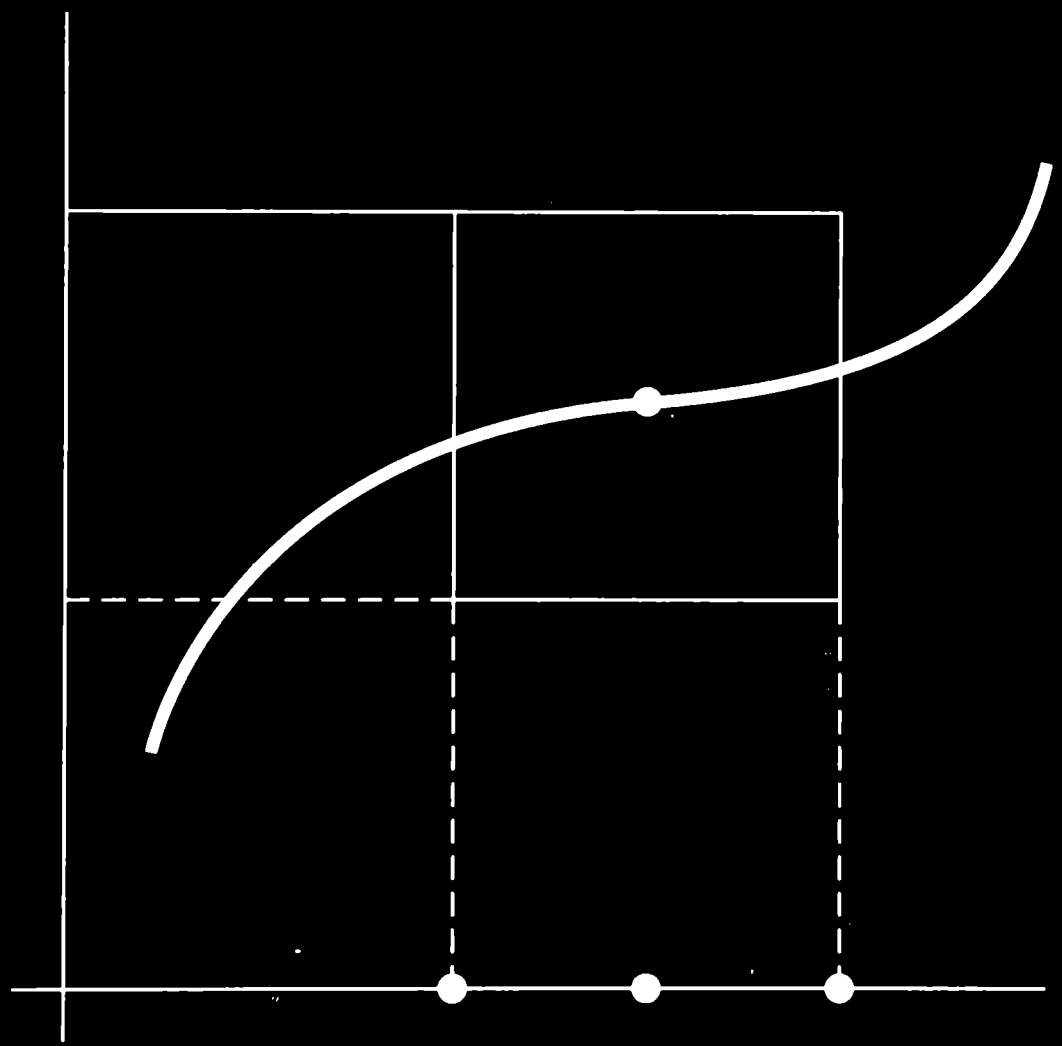
Cours d'Analyse

1 Analyse vectorielle

Srishti D. Chatterji



Presses polytechniques et universitaires romandes

Cours d'Analyse

1 Analyse vectorielle

Cours d'Analyse

1 Analyse vectorielle

Srishti D. Chatterji

Presses polytechniques et universitaires romandes

DANS LA COLLECTION «MATHÉMATIQUES»

Calcul différentiel et intégral
Jacques Douchet et Bruno Zwahlen
1 Fonctions réelles d'une variable réelle
2 Fonctions réelles de plusieurs variables réelles – Exercices résolus
3 Fonctions réelles d'une variable réelle – Exercices résolus
4 Fonctions réelles de plusieurs variables réelles – Exercices résolus

Analyse non standard
Alain Robert

Algèbre linéaire
Renzo Cairoli

Eléments de programmation linéaire avec applications aux graphes
Dominique de Werra

Introduction à la logique
André Delessert

Initiation aux probabilités
Sheldon M. Ross

DANS LA COLLECTION «CAHIERS MATHÉMATIQUES DE L'EPFL»

Mécanique analytique 1 et 2
Philippe Choquard

Introduction au calcul des variations
Bernard Dacorogna

Composition de polytopes combinatoires
François Margot

Les Presses polytechniques et universitaires romandes sont une fondation scientifique dont le but est principalement la diffusion des travaux de l'Ecole polytechnique fédérale de Lausanne, d'autres universités francophones ainsi que des écoles techniques supérieures. Le catalogue de leurs publications peut être obtenu par courrier aux Presses polytechniques et universitaires romandes, EPFL – Centre Midi, CH-1015 Lausanne, par E-Mail à ppur@epfl.ch, par téléphone au (0)21 693 41 40, ou par fax au (0)21 693 40 27.

Vous pouvez consulter notre catalogue général sur notre serveur Internet http://sente.epfl.ch/webs/ppur/

ISBN 2-88074-314-1
© 1997, Presses polytechniques et universitaires romandes,
CH – 1015 Lausanne
Tous droits réservés.
Reproduction, même partielle, sous quelque forme ou sur quelque support que ce soit, interdite sans l'accord écrit de l'éditeur.

à la mémoire de ma mère
à mon père

Avant-propos

Le cours d'analyse de deuxième année pour les futurs ingénieurs mathématiciens et physiciens à l'Ecole Polytechnique Fédérale de Lausanne (EPFL) comporte trois parties :

1. Analyse vectorielle
2. Analyse complexe
3. Equations différentielles ordinaires et aux dérivées partielles.

Le présent volume concerne l'analyse vectorielle ; il est prévu de publier deux autres volumes pour les deux autres parties du cours. L'objectif de ce volume et de la première partie du cours est de présenter la théorie mathématique autour des théorèmes intégraux qui sont souvent associés aux noms de Gauss, Green et Stokes.

L'ouvrage s'adresse aux lecteurs qui ont une connaissance de l'algèbre linéaire et du calcul différentiel et intégral dans $\mathbb{R}^N$ comme enseignés dans beaucoup d'universités européennes. Les premiers trois chapitres récapitulent le calcul différentiel dans $\mathbb{R}^N$; après une présentation des considérations topologiques essentielles dans $\mathbb{R}^N$, on reprend les propriétés des fonctions dérivables, se bornant généralement aux fonctions de classe C^k, $k \geqslant 1$. Le chapitre quatre explique les théorèmes d'inversion locale, des fonctions implicites et du rang constant, un des objectifs principaux étant l'introduction des sous-variétés différentielles de $\mathbb{R}^N$. Le chapitre suivant présente la théorie de l'intégration de Lebesgue dans $\mathbb{R}^N$ en guise de théorie de l'intégration dans un espace abstrait ; en évitant de discuter en détail les problèmes associés à la notion de mesurabilité et en admettant un petit nombre d'énoncés clairement formulés, on a pu exposer les théorèmes centraux de la théorie comme le théorème de convergence dominée et le théorème de Fubini ; le théorème de changement de variables dans les intégrales multiples est démontré en détail dans une formulation utile pour les applications. L'analyse vectorielle est traitée dans le chapitre six, essentiellement dans $\mathbb{R}^2$ et $\mathbb{R}^3$; les intégrales curvilignes et superficielles sont introduites et les propriétés usuelles des opérateurs grad, div, rot sont étudiées ; les théorèmes principaux, appelés par la suite théorèmes de Green (dans $\mathbb{R}^2$), Gauss et Stokes (dans $\mathbb{R}^3$), sont énoncés précisément ; mais, dans ce chapitre, on se contente de les vérifier dans certains cas simples et illustratifs. L'objectif de ce chapitre est double : donner les outils nécessaires pour les cours de physique (par exemple la dynamique des fluides ou l'électromagnétisme) et préparer le terrain pour le théorème de Stokes dans $\mathbb{R}^N$. Le dernier chapitre est presque entièrement consacré à la formulation et à la démonstration d'une version du théorème de Stokes dans $\mathbb{R}^N$; pour plus de clarté, le théorème de la divergence est formulé et démontré séparément avant

de passer au théorème de Stokes général. Ces théorèmes sont énoncés dans une version qui admet certains domaines anguleux qui apparaissent naturellement dans les applications envisagées au chapitre six.

Les chapitres six et sept forment une unité ; néanmoins, il est possible de les lire indépendamment. D'ailleurs, en général, les chapitres et mêmes les différentes sections d'un chapitre sont composés de telle façon que les lecteurs débutants puissent les lire de manière presque autonome.

Etant donné les lecteurs auxquels l'ouvrage est destiné, la présentation est généralement très détaillée ce qui explique le volume de ce livre malgré son étendue mathématique limitée. Naturellement, je suis conscient que les débutants trouveront le dernier chapitre difficile et que certaines parties des chapitres quatre et cinq ne seront pas faciles pour eux. Les théorèmes autour de l'inversion locale ont toujours présenté des difficultés et l'intégration multidimensionnelle a ses complications inévitables. J'ai décidé d'affronter ces problèmes sans ambages ; le théorème de Stokes et ses préparatifs forment une première étape importante dans l'apprentissage de l'analyse sur les variétés et un effort sérieux est indispensable pour la franchir. En écrivant le chapitre sept j'ai essayé de rendre la tâche aussi abordable que possible.

Etant donné l'objectif limité du livre, j'ai généralement évité la discussion des questions, fussent-elles importantes, qui ne sont pas directement liées aux développements des chapitres six et sept. Mais, ici et là, j'ai ajouté quelques commentaires et notes historiques pour montrer quelques-uns des multiples aspects de la théorie présentée. Les remarques historiques sont extrêmement sommaires et elles n'ont pas l'ambition de fixer la vérité historique autour du sujet ; tout au plus, elles peuvent servir comme guide pour une première réflexion sur le développement historique de la matière.

Presque chaque section contient une série d'exercices ; ceux-ci sont de deux types. Les uns demandent un calcul direct basé sur la théorie développée précédemment ; pour ces exercices on a fourni les réponses à la fin du livre. Les autres contiennent des exemples et des contre-exemples illustrant la théorie ; ces exercices sont essentiellement résolus par une série de suggestions intercalées à l'intérieur même de leur formulation.

Dans mon cours tel qu'il est professé à l'EPFL, je dispose approximativement de 26 semaines (3 heures de cours et 2 heures d'exercices hebdomadaires) pour couvrir les trois parties signalées du cours ; seulement 6 semaines sont consacrées à la matière du présent livre. Dans ce laps de temps, j'arrive à présenter une partie importante du chapitre six et effleurer le formalisme des formes différentielles après avoir établi la terminologie concernant la topologie de $\mathbb{R}^N$ et la dérivabilité des fonctions de plusieurs variables réelles. Puisque je n'ai pas le temps de développer l'appareil théorique autour du théorème de Stokes, j'ai décidé de le présenter ici en détail pour ceux qui désireraient s'informer sur le sujet en autodidactes. On peut imaginer un cours de deuxième cycle introductif basé sur le chapitre sept et ses préparatifs ; d'autres utilisations de la matière présentée ici sont aussi envisageables.

Quelques pages préliminaires fixent les conventions et les notations utilisées ; on conseille aux lecteurs débutants de les parcourir rapidement avant d'aborder le texte ; quelques rappels, sans développement détaillé, sont inclus pour leur commodité.

En guise de conclusion, j'ai ajouté quelques lignes à la fin du livre comme suggestions pour l'approfondissement de la matière traitée ici.

S. D. Chatterji
Lausanne

Table des matières

AVANT-PROPOS vii

CONVENTIONS, NOTATIONS ET RAPPELS 1

1 Ensembles et fonctions 1
2 Nombres réels 3
3 L'espace $\mathbb{R}^N$ $(N = 1, 2, \ldots)$ 5
4 Permutations 10
5 Déterminants 11
6 Espace $\mathbb{R}^3$ 13
7 Cardinalité 14
8 Quelques fonctions réelles 14
9 Conventions diverses 15

CHAPITRE 1 TOPOLOGIE DE $\mathbb{R}^N$

1.1 Préliminaires 18
1.1.1 Définitions 18
1.1.2 Exercices 20
1.2 Ensembles ouverts 21
1.2.1 Proposition 21
1.2.2 Exercices 23
1.3 Terminologie topologique 24
1.3.1 Quelques relations entre A, $\mathring{A}$, $\overline{A}$, $\mathrm{Fr}\,A$, $\mathrm{ext}\,A$ .. 24
1.3.2 Proposition 25
1.3.3 Proposition 25
1.3.4 Remarque 26
1.3.5 Exercices 27
1.4 Limites 28
1.4.1 Limites des suites 28
1.4.2 Proposition 29
1.4.3 Séries 30
1.4.4 Proposition 30
1.4.5 Remarque 31
1.4.6 Limites des fonctions 32
1.4.7 Exercices 34

1.5 Topologie induite 35
1.5.1 Définitions 35
1.5.2 Remarques 37
1.5.3 Exercices 37
1.6 Fonctions continues 38
1.6.1 Définitions 38
1.6.2 Proposition 39
1.6.3 Remarques 41
1.6.4 Exercices 42
1.7 Compacité 44
1.7.1 Proposition 44
1.7.2 Remarques 48
1.7.3 Proposition 50
1.7.4 Exercices 52
1.8 Connexité 53
1.8.1 Définitions 53
1.8.2 Proposition 54
1.8.3 Remarques sur la connexité 57
1.8.4 Exercices 59
1.9 Remarques supplémentaires sur les fonctions continues 60
1.9.1 La fonction $x \mapsto d(x, A)$ 60
1.9.2 Limites des fonctions continues 63
1.9.3 Prolongement des fonctions continues 65
1.9.4 Fonctions convexes 67
1.9.5 Un théorème du point fixe 70
1.9.6 Exercices 71
1.10 Exercices 72
1.11 Commentaires 74

CHAPITRE 2 DÉRIVABILITÉ

2.1 La dérivée 77
2.1.1 Généralités 77
2.1.2 Proposition 79
2.1.3 Remarques 81
2.1.4 Proposition 83
2.1.5 Remarques 84
2.1.6 Exercices 85

2.2 Dérivation des fonctions composées 88
2.2.1 Proposition (*formule de dérivation des fonctions composées*) 88
2.2.2 Remarques 89
2.2.3 Manipulation des différentielles 90
2.2.4 Exercices 91
2.3 Formule des accroissements finis 93
2.3.1 Proposition 94
2.3.2 Corollaire 95
2.3.3 Corollaire 96
2.3.4 Remarques 96
2.3.5 Exercices 99
2.4 Maximum et minimum 100
2.4.1 Proposition (*condition nécessaire de premier ordre pour un extremum relatif*) 100
2.4.2 Remarques 101
2.4.3 Un problème d'extremum absolu dans $\mathbb{R}^N$ ... 102
2.4.4 Un problème d'extremum absolu dans $\mathbb{R}^2$ 104
2.4.5 Exercices 108
2.5 Convergence des dérivées 108
2.5.1 Proposition 110
2.5.2 Corollaire 111
2.5.3 Remarques 112
2.6 Interprétations de la dérivée 113
2.7 Exercices 116

CHAPITRE 3 DÉRIVÉES D'ORDRE SUPÉRIEUR
3.1 Définitions et notations principales 121
3.1.1 Généralités 121
3.1.2 Remarques 123
3.1.3 Certains opérateurs différentiels 123
3.2 Dérivées mixtes 126
3.2.1 Proposition 126
3.2.2 Proposition (*théorème de Schwarz*) 126
3.2.3 Remarque 128
3.2.4 Proposition 128
3.2.5 Remarques et exemples 128
3.2.6 Exercices 129

3.3 Formule de Taylor 132
3.3.1 Notation 132
3.3.2 Proposition (*formule de Taylor : développement limité d'ordre n*) 134
3.3.3 Remarques 136
3.3.4 Généralisation aux fonctions vectorielles 136
3.3.5 Exercices 137
3.4 Maximum et minimum 138
3.4.1 Matrices de type positif ou négatif 139
3.4.2 Lemme 140
3.4.3 Remarque 141
3.4.4 Proposition 141
3.4.5 Remarque 143
3.4.6 Une méthode pour la vérification de la positivité d'une matrice 143
3.4.7 Cas particulier de deux variables 146
3.4.8 Exercices 149
3.5 Coordonnées polaires dans $\mathbb{R}^2$ 149
3.5.1 Introduction 149
3.5.2 Proposition 150
3.5.3 Remarques 151
3.5.4 Exercices 152
3.6 Fonctions harmoniques 153
3.6.1 Proposition (*principe du maximum faible*) 153
3.6.2 Corollaire (*principe du minimum faible*) 154
3.6.3 Remarques 154
3.6.4 Unicité de la solution du problème de Dirichlet 155
3.7 Fonctions convexes 156
3.7.1 Proposition 156
3.7.2 Remarques 157
3.7.3 Proposition 157
3.7.4 Remarques 159
3.7.5 Inégalité de Jensen 162
3.7.6 Inégalité de Minkowski 162
3.7.7 Remarques 163
3.7.8 Inégalité de Hölder 164
3.7.9 Exercices 164

3.8 Le laplacien en différentes coordonnées 167
3.8.1 Introduction 167
3.8.2 Le laplacien dans $\mathbb{R}^2$ en coordonnées polaires 168
3.8.3 Le laplacien dans $\mathbb{R}^3$ en coordonnées sphériques 168
3.8.4 Inversion, transformée de Kelvin 170
3.9 Quelques définitions générales 172
3.9.1 Définitions de $C^k(A;\mathbb{R}^M)$ 172
3.9.2 Difféomorphismes 173
3.9.3 Proposition 174
3.9.4 Exemples de difféomorphismes 175
3.10 Quelques remarques sur les dérivées d'ordre supérieur 176
3.11 Exercices 178
3.12 Commentaires 179

CHAPITRE 4 FONCTIONS IMPLICITES
4.1 Préliminaires 181
4.1.1 Position du problème 181
4.1.2 Solution plausible du problème 182
4.1.3 Calcul de la dérivée des fonctions implicites .. 183
4.1.4 Exercices 185
4.2 Fonctions inverses 186
4.2.1 Terminologie 186
4.2.2 Proposition (*théorème des fonctions inverses ou théorème d'inversion locale*) 186
4.2.3 Lemme (*injectivité*) 186
4.2.4 Lemme (*application ouverte*) 187
4.2.5 Démonstration de la proposition 4.2.2 188
4.2.6 Proposition 189
4.2.7 Remarques 189
4.2.8 Les images 191
4.2.9 Théorèmes d'inversion globale 192
4.2.10 Exercices 194
4.3 Théorème des fonctions implicites 196
4.3.1 Proposition (*théorème des fonctions implicites*) 196
4.3.2 Remarques 198
4.3.3 Exercices 199

4.4 Théorème du rang constant 200
4.4.1 Introduction 200
4.4.2 Proposition (*théorème du rang constant*) 201
4.4.3 Remarques et définitions 204
4.4.4 Proposition 205
4.4.5 Immersions 208
4.4.6 Submersions 208
4.4.7 Valeurs régulières 209
4.4.8 Applications de rang r 210
4.4.9 Dépendance fonctionnelle 211
4.4.10 Exercices 213
4.5 Extremums liés 215
4.5.1 Introduction 215
4.5.2 Méthode des multiplicateurs de Lagrange 217
4.5.3 Optimisation non linéaire 218
4.5.4 Exercices 221
4.6 Remarques 224

CHAPITRE 5 INTÉGRATION

5.1 Remarques préliminaires 228
5.1.1 L'intégrale de Riemann : rappels intuitifs 228
5.1.2 Intégrales multiples : rappels intuitifs 231
5.1.3 Nécessité d'une théorie de l'intégration générale 231
5.2 Intégrale de Lebesgue 232
5.2.1 Espaces boréliens 233
5.2.2 Mesures 234
5.2.3 Fonctions mesurables 236
5.2.4 Intégrales des fonctions positives 238
5.2.5 Proposition 239
5.2.6 Proposition 239
5.2.7 Intégrales des fonctions non nécessairement positives 242
5.2.8 Proposition 243
5.2.9 Proposition 246
5.2.10 Intégrale comme la limite d'une somme 247
5.2.11 Ensembles négligeables, fonctions μ-mesurables 248

5.2.12 Proposition 249
5.2.13 Intégrales de Lebesgue dans $\mathbb{R}^N$ 250
5.2.14 Proposition 251
5.2.15 Exercices 252
5.3 Intégrales qui dépendent d'un paramètre 257
5.3.1 Proposition 258
5.3.2 Corollaire 258
5.3.3 Proposition (*dérivation sous le signe d'intégration*) 259
5.3.4 Corollaire 260
5.3.5 Exercices 261
5.4 Théorème de Fubini 264
5.4.1 Définitions préliminaires 264
5.4.2 Proposition (*existence de la mesure produit*) . 265
5.4.3 Proposition (théorème de Fubini) 266
5.4.4 Cas de $\mathbb{R}^N$ 267
5.4.5 Exercices 270
5.5 Mesures dans $\mathbb{R}^N$ 276
5.5.1 Mesures de Radon 276
5.5.2 Proposition 278
5.5.3 Remarque 279
5.6 Mesure de Lebesgue dans $\mathbb{R}^N$ 280
5.6.1 Généralités 280
5.6.2 Proposition 281
5.6.3 Transformations affines 281
5.6.4 Proposition 282
5.6.5 Volume d'un parallélépipède dans $\mathbb{R}^N$ 283
5.6.6 Proposition 285
5.6.7 Ensembles négligeables dans $\mathbb{R}^N$ 288
5.7 Changement de variables dans les intégrales 290
5.7.1 Introduction 290
5.7.2 Proposition 291
5.7.3 Remarques 293
5.7.4 Formulation abstraite de la formule de changement de variables dans les intégrales .. 294
5.8 Quelques cas importants de changement de variables dans les intégrales multiples 296
5.8.1 Transformation affine 296

5.8.2 Transformation polaire dans $\mathbb{R}^2$ 296
5.8.3 Transformation cylindrique dans $\mathbb{R}^3$ 297
5.8.4 Transformation sphérique dans $\mathbb{R}^3$ 297
5.8.5 Transformation sphérique dans $\mathbb{R}^N$ 298
5.9 Intégrales des fonctions vectorielles 300
5.10 Exercices 304
5.11 Approximation par des fonctions de classe C^∞ 313
5.11.1 Généralités 313
5.11.2 Lemme 315
5.11.3 Lemme 316
5.11.4 Proposition 317
5.11.5 Proposition 318
5.11.6 Proposition (*théorème de Weierstrass*) 318
5.11.7 Remarques 321
5.12 Quelques compléments 321
5.12.1 Inégalité de Jensen 321
5.12.2 Inégalités de Hölder et de Minkowski 323
5.12.3 Formule d'Euler-Maclaurin, formule de Stirling 325
5.12.4 Exercices 328
5.13 Commentaires 329
5.14 Appendice 331

CHAPITRE 6 ANALYSE VECTORIELLE

6.1 Intégrales curvilignes dans $\mathbb{R}^N$ 333
6.1.1 Définitions 333
6.1.2 Terminologie 334
6.1.3 Interprétation intuitive des intégrales curvilignes 335
6.1.4 Proposition 336
6.1.5 Explications au sujet de la proposition 6.1.4 .. 337
6.1.6 Représentation d'un lacet 338
6.1.7 Paramétrage naturel 339
6.1.8 Proposition 341
6.1.9 Champ dérivant d'un potentiel 341
6.1.10 Notations et conventions 343
6.1.11 Proposition 344
6.1.12 Compléments à la proposition 6.1.11 346

6.1.13 Explication physique du terme conservatif.... 346
6.1.14 Ensemble étoilé..... 347
6.1.15 Proposition..... 348
6.1.16 Remarques..... 350
6.1.17 Courbes..... 351
6.1.18 Exercices..... 352
6.2 Théorème de Green..... 357
6.2.1 Explications intuitives..... 357
6.2.2 Une vérification..... 358
6.2.3 Théorème de Jordan..... 360
6.2.4 Domaines réguliers dans $\mathbb{R}^2$ et orientation de leurs frontières..... 361
6.2.5 Proposition (*formulation précise du théorème de Green*)..... 363
6.2.6 Calcul d'aire..... 364
6.2.7 Formule de changement de variables dans $\mathbb{R}^2$. 364
6.2.8 Le flux à travers une courbe plane..... 366
6.2.9 Remarques sur la validité du théorème de Green..... 367
6.2.10 Une condition géométrique..... 368
6.2.11 Exercices..... 370
6.3 Intégrales de surface dans $\mathbb{R}^3$..... 373
6.3.1 Morceaux de surfaces régulières dans $\mathbb{R}^3$..... 373
6.3.2 Remarques concernant la propriété homéomorphique de ψ..... 375
6.3.3 Intégrales de surface..... 376
6.3.4 Normale unitaire, orientation, flux..... 378
6.3.5 Changement de paramétrage..... 380
6.3.6 D'autres notations pour le flux..... 381
6.3.7 Intégrales de surface : définition globale..... 383
6.3.8 Flux à travers une surface : définition globale..... 384
6.3.9 Exemples de surfaces orientables..... 385
6.3.10 Remarques géométriques sur la définition de l'aire d'une surface..... 387
6.3.11 Sur la définition d'une surface..... 388
6.3.12 Exercices..... 389

6.4 Théorème de Gauss dans $\mathbb{R}^3$ 397
6.4.1 Domaines réguliers dans $\mathbb{R}^3$ 397
6.4.2 Proposition (*théorème de la divergence*) 398
6.4.3 Une vérification du théorème de la divergence 399
6.4.4 Angle solide 400
6.4.5 Loi de Gauss 402
6.4.6 Formule de Green 403
6.4.7 Propriétés des fonctions harmoniques 404
6.4.8 Poussée d'Archimède 406
6.4.9 Exercices 408
6.5 Théorème de Stokes dans $\mathbb{R}^3$ 410
6.5.1 Introduction 410
6.5.2 Domaines réguliers et leurs bords dans une surface 411
6.5.3 Orientation du bord 412
6.5.4 Proposition (*théorème de Stokes dans* $\mathbb{R}^3$) ... 413
6.5.5 Une vérification du théorème de Stokes 414
6.5.6 Remarques 416
6.5.7 Exercices 417
6.6 Quelques problèmes concernant les champs vectoriels dans $\mathbb{R}^3$ 419
6.6.1 Formules concernant les champs vectoriels ... 419
6.6.2 Champ irrotationnel, champ incompressible .. 421
6.6.3 Proposition 422
6.6.4 Remarques 424
6.6.5 Proposition (*théorème de Helmholtz*) 425
6.6.6 Remarques 425
6.6.7 Exercices 426
6.7 Equations de Maxwell 430
6.7.1 Introduction 430
6.7.2 Interdépendance des équations (6.38) 431
6.7.3 Lois de conservation 431
6.7.4 Circulations de E et H 432
6.7.5 Vecteur de Poynting 432
6.7.6 Solutions des équations 433
6.8 Coordonnées curvilignes dans $\mathbb{R}^3$ 435
6.8.1 Introduction 435
6.8.2 Notations et conventions supplémentaires 436

6.8.3 Proposition 437
6.8.4 Esquisse d'une autre démonstration 440
6.8.5 Exercices 442
6.9 Commentaires 443

CHAPITRE 7 THÉORÈME DE STOKES GÉNÉRALISÉ
7.1 Formalisme des formes différentielles 445
7.1.1 Introduction 445
7.1.2 Symbolisme des formes différentielles extérieures 446
7.1.3 Différentiation extérieure des formes 449
7.1.4 Formes fermées, formes exactes 451
7.1.5 Changement de variables dans les formes différentielles 452
7.1.6 Intégration des formes différentielles : esquisse 453
7.1.7 Exercices 455
7.2 Partitions de l'unité 457
7.2.1 Préliminaires 457
7.2.2 Proposition 458
7.2.3 Remarques 460
7.2.4 Proposition 461
7.2.5 Remarques 461
7.2.6 Proposition 462
7.2.7 Proposition 463
7.2.8 Proposition 464
7.2.9 Remarques 465
7.2.10 Exercices 466
7.3 Intégration dans les variétés différentielles 467
7.3.1 Préliminaires 467
7.3.2 Changement de paramétrage local 469
7.3.3 Espaces tangents 470
7.3.4 Volume d'un parallélépipède 472
7.3.5 Fomule de Binet-Cauchy 475
7.3.6 L'élément de volume 476
7.3.7 Mesure naturelle dans une variété 478
7.3.8 Ensembles négligeables dans une variété 481
7.3.9 Ensembles k-négligeables 482
7.3.10 Exercices 483

7.4 Théorème de la divergence dans $\mathbb{R}^N$ 487
7.4.1 Introduction 487
7.4.2 Frontière régulière 487
7.4.3 Normale dirigée vers l'extérieur 490
7.4.4 Proposition 492
7.4.5 Proposition 493
7.4.6 Proposition (*théorème de la divergence dans* $\mathbb{R}^N$) 495
7.4.7 Remarques 499
7.4.8 Exercices 500
7.5 Formes multilinéaires alternées 503
7.5.1 Introduction 503
7.5.2 Préliminaires 503
7.5.3 Produit extérieur 504
7.5.4 Proposition 507
7.5.5 Remarques 507
7.5.6 Compléments 510
7.5.7 Exercices 512
7.6 Formes différentielles 513
7.6.1 Introduction 513
7.6.2 Définition des formes différentielles 513
7.6.3 Produit extérieur des formes différentielles ... 514
7.6.4 Différentiation extérieure 515
7.6.5 Changement de variables 517
7.7 Orientation 521
7.7.1 Variété différentielle orientée 521
7.7.2 Orientation des espaces vectoriels 522
7.7.3 Orientation des espaces tangents 524
7.7.4 Orientation des hypersurfaces 526
7.7.5 Proposition 527
7.7.6 Orientation d'un bord 529
7.7.7 Compléments 531
7.8 Intégration des formes différentielles 532
7.8.1 Introduction 532
7.8.2 Intégration d'une N-forme 532
7.8.3 Intégration d'une k-forme 534
7.8.4 Une inégalité pour les intégrales d'une k-forme 536

7.8.5 Intégration sur une hypersurface orientée 538
7.8.6 Remarques 540
7.8.7 Une formule utile 540
7.8.8 Un théorème de convergence 542
7.9 Théorème de Stokes généralisé 544
7.9.1 Proposition (*théorème de Stokes généralisé*).. 544
7.9.2 Corollaire 544
7.9.3 Remarques 548
7.9.4 Version simpliciale du théorème de Stokes.... 549
7.9.5 Théorème de Stokes pour les variétés à bord . 551
7.9.6 Exercices 553
7.10 Formes fermées, formes exactes : retour 554
7.10.1 Introduction 554
7.10.2 Homotopie des applications 554
7.10.3 Théorème de Poincaré 556
7.10.4 Proposition (*théorème de Poincaré*) 556
7.10.5 Remarques 557
7.10.6 Les 1-formes 559
7.10.7 Homotopie des chemins 560
7.10.8 Proposition 561
7.10.9 Connexité simple 563
7.10.10 Proposition 564
7.10.11 Remarques 564
7.10.12 Exercices 566

CONCLUSION 567

RÉPONSES AUX EXERCICES 569

BIBLIOGRAPHIE 583

INDEX 585

Conventions, notations et rappels

1 Ensembles et fonctions

$x \in A$ x appartient à l'ensemble A

$x \notin A$ x n'appartient pas à l'ensemble A

$A \subset B$ l'ensemble A est contenu dans l'ensemble B ; ceci équivaut à $B \supset A$ et n'exclut pas que $A = B$; on dit que A est une partie de B

A^c le complémentaire de A

$A \cup B$ la réunion des ensembles A et B

$A \cap B$ l'intersection des ensembles A et B

$A \setminus B$ l'ensemble formé des éléments de A qui ne sont pas dans B ; donc $A \setminus B = A \cap B^c$

$\emptyset$ l'ensemble vide

Une ***fonction*** (les mots fonction, ***application***, ***correspondance***, ***transformation*** sont des synonymes) est notée comme $f : X \to Y$ où X est le ***domaine*** de définition de f et Y, l'***espace image***, contient les ***images*** ou les ***valeurs*** $f(x)$, $x \in X$; on écrit aussi $x \mapsto f(x)$, $x \in X$, pour indiquer la fonction f ; pour $x \in X$, il existe un unique élément $y \in Y$ tel que $y = f(x)$; f s'appelle ***injective*** si $x \neq x'$ dans X entraîne que $f(x) \neq f(x')$ dans Y ; f s'appelle ***surjective*** si chaque $y \in Y$ est de la forme $y = f(x)$ pour $x \in X$; f est dite ***bijective*** si f est à la fois injective et surjective. Si $f : X \to Y$ on utilise aussi les phrases du type : f est une fonction sur X à valeurs dans Y.

Si $f : I \to J$ est une fonction, on dit que $\{f(i)\}_{i \in I}$ est une ***famille d'éléments dans J indexée pour I*** ; on note cette famille aussi par $\{f_i\}_{i \in I}$ ou simplement $\{f_i\}$ si I n'est pas en doute.

Soit $f : I \to J$; il est souvent commode de décrire l'ensemble $f(I) \subset J$ comme suit :

$$f(I) = \{f(i) : i \in I\}.$$

Il est important de distinguer l'ensemble $f(I) \subset J$ de la famille $\{f(i)\}_{i \in I}$ dans J.

Une famille finie (appelée aussi ***suite finie***) dans un ensemble X est notée $\{x_i\}_{i=1,\ldots,n}$; ici, les x_i ne sont pas nécessairement distincts ; par contre, une partie finie de X est aussi notée $\{x_i\}_{i=1,\ldots,n}$ ou $\{x_1, \ldots, x_n\}$ mais avec la convention tacite que les x_i sont *distincts*.

Si $\{A_i\}_{i \in I}$ est une famille de parties d'un ensemble X, on note par

$$\bigcup_{i \in I} A_i \quad \text{ou} \quad \bigcup \{A_i : i \in I\} \quad \text{ou} \quad \bigcup_i A_i$$

la réunion des A_i ; on note par

$$\bigcup_{i\in I} A_i \quad \text{ou} \quad \bigcup\{A_i : i \in I\} \quad \text{ou} \quad \bigcup_i A_i$$

l'intersection des A_i.

Rappelons que les opérations ensemblistes $\bigcap, \bigcup$ sont commutatives, associatives et distributives. On a les formules suivantes :

$$A \cap \bigcup_{i\in I} A_i = \bigcup_{i\in I} (A \cap A_i)$$

$$A \cup \bigcap_{i\in I} A_i = \bigcap_{i\in I} (A \cup A_i)$$

$$\left\{\bigcup_{i\in I} A_i\right\}^c = \bigcap_{i\in I} A_i^c, \quad \left\{\bigcap_{i\in I} A_i\right\}^c = \bigcup_{i\in I} A_i^c.$$

Si $X_1, \ldots, X_n$ sont n ensembles, on note par $X_1 \times \cdots \times X_n$ l'ensemble formé de tous les n-uples $(x_1, \ldots, x_n)$ avec $x_i \in X_i$, $i = 1, \ldots, n$; on écrit

$$X_1 \times \cdots \times X_n = \{(x_1, \ldots, x_n) : x_i \in X_i,\ i = 1, \ldots, n\} ;$$

$X_1 \times \cdots \times X_n$ s'appelle le ***produit cartésien*** des X_i, $i = 1, \ldots, n$.

Soit $f : X \to Y$ une fonction ; l'ensemble

$$G_f = \{(x, f(x)) : x \in X\} \subset X \times Y$$

s'appelle le ***graphe*** de f ; si $A \subset X$, on écrit

$$f(A) = \{f(x) : x \in A\} \subset Y ;$$

$f(A)$ s'appelle l'***image*** de A sous f ; si $A \subset Y$, on écrit

$$f^{-1}(A) = \{x \in X : f(x) \in A\} \subset X ;$$

$f^{-1}(A)$ s'appelle l'***image inverse*** de A sous f.

Le symbole f^{-1} est utilisé dans un autre contexte. Soit $f : X \to Y$ une fonction *bijective* ; pour tout $y \in Y$ il existe alors un unique $x \in X$ tel que $f(x) = y$; on écrit $f^{-1}(y) = x$; alors $f^{-1} : Y \to X$ est une fonction telle que

$$f^{-1}(f(x)) = x,\ x \in X ; \quad f(f^{-1}(y)) = y,\ y \in Y ;$$

f^{-1} s'appelle la fonction ***inverse*** ou ***réciproque*** de f.

Soit $f : X \to Y$; si $A \subset X$ on écrit $f|_A$ pour la ***restriction*** de f à A ; donc $f|_A : A \to Y$ est une fonction telle que $f(x) = f|_A(x)$ si $x \in A$; s'il n'y a aucune confusion à craindre, nous indiquerons $f|_A$ par $f : A \to Y$.

Soient $f : X \to Y$, $g : Y \to Z$ deux fonctions ; alors $g \circ f : X \to Z$ est la fonction définie par

$$(g \circ f)(x) = g\big(f(x)\big), \quad x \in X ;$$

$g \circ f$ s'appelle la ***composition*** des fonctions f et g.

Soit X un ensemble ; une ***partition*** de X est une famille de sous-ensembles $\{A_i\}_{i \in I}$ de X telle que les A_i sont deux à deux disjoints et leur réunion est égale à X ; la partition s'appelle ***finie*** (ou ***dénombrable***, cf. § 7) si I est fini ou dénombrable.

2 Nombres réels

Nous notons par $\mathbb{R}$ le corps des nombres réels. Rappelons que $\mathbb{R}$ est caractérisé par la description suivante : $\mathbb{R}$ est un corps ***commutatif***, ***ordonné***, ***archimédien*** et ***complet***. Expliquons les deux derniers termes ; dans $\mathbb{R}$, on note par $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ les sous-ensembles suivants :

$$\mathbb{N} = \{0, 1, 2, \ldots\} \quad \text{nombres naturels}$$

$$\mathbb{Z} = \{0, 1, -1, 2, -2, \ldots\} \quad \text{nombres entiers}$$

$$\mathbb{Q} = \left\{\frac{m}{n} : m \in \mathbb{Z}\,,\ n \in \mathbb{N},\ n \geqslant 1\right\} \quad \text{nombres rationnels.}$$

$\mathbb{R}$ est archimédien : pour tout $x > 0$, $y > 0$, il existe $n \in \mathbb{N}$ tel que $nx > y$. $\mathbb{R}$ est complet : toute suite de Cauchy dans $\mathbb{R}$ converge vers un élément de $\mathbb{R}$. Nous supposons connues les définitions et les propriétés des suites des nombres réels invoquées dans la phrase précédente.

Si $A \subset \mathbb{R}$, A^* sera l'ensemble $A \setminus \{0\}$ et A_+ sera l'ensemble $\{x \in A : x \geqslant 0\}$. Ainsi,

$$\mathbb{Z}_+ = \mathbb{N}, \quad \mathbb{N}^* = \{1, 2, \ldots\}\,.$$

De même, $A_- = \{x \in A : x \leqslant 0\}$.

Par $\overline{\mathbb{R}}$ nous noterons l'ensemble $\mathbb{R} \cup \{\infty\,, -\infty\}$ où ∞, $-\infty$ sont deux symboles distincts et différents des élements de $\mathbb{R}$; parfois, ∞ sera noté par $+\infty$. On écrit : $-\infty < x < \infty$ pour tout $x \in \mathbb{R}$. Si A est une partie non vide de $\overline{\mathbb{R}}$, on définit

$$M = \sup\big\{x : x \in A\big\} = \sup_{x \in A} x$$

comme l'unique élément $M \in \overline{\mathbb{R}}$ tel que

$$M \geqslant x\,,\ x \in A\,; \quad \text{si } M' \geqslant x\,,\ x \in A \text{ alors } M' \geqslant M$$

(M est le plus petit majorant des éléments de A) ; de manière analogue, on définit

$$m = \inf\{x : x \in A\} = \inf_{x \in A} x$$

comme l'unique élément $m \in \overline{\mathbb{R}}$ tel que

$$m \leqslant x\,,\ x \in A\,; \quad \text{si } m' \leqslant x\,,\ x \in A \text{ alors } m' \leqslant m$$

(m est le plus grand minorant des éléments de A).

Si X est un ensemble quelconque, une ***suite*** dans X est une famille d'éléments de X indexée par $\mathbb{N}$; une suite sera notée comme

$$\{x_n\}_{n\in\mathbb{N}} \quad \text{ou} \quad \{x_n\}_{n\geqslant 0} \quad \text{ou} \quad \{x_n\}_{n=0,1,\ldots} \quad \text{ou simplement } \{x_n\}\,;$$

souvent, les familles indexées par $\{k,\ k+1,\ \ldots\}$ où $k \in \mathbb{Z}$ seront aussi appelées les suites ; elles seront notées

$$\{x_n\}_{n\geqslant k} \quad \text{ou} \quad \{x_n\}_{n=k,k+1,\ldots} \quad \text{ou simplement} \quad \{x_n\} \text{ ou } \{x_i\} \text{ etc.}$$

Nous écrirons simplement $\{x_n\}$ ou $\{x_i\}$, etc., s'il n'y a aucune confusion à craindre concernant le domaine de variation de n ou i.

Soit $\{x_n\}_{n\in\mathbb{N}}$ une suite dans $\overline{\mathbb{R}}$. On pose

$$\limsup_{n\to\infty} x_n = \inf_{n\geqslant 0}\left(\sup_{k\geqslant n} x_k\right)$$

$$\liminf_{n\to\infty} x_n = \sup_{n\geqslant 0}\left(\inf_{k\geqslant n} x_k\right)\,;$$

si $M_n = \sup_{k\geqslant n} x_k$, on voit que

$$M_n \geqslant M_{n+1}\,, \quad n \geqslant 0\,;$$

donc

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} M_n = \lim_{n\to\infty}\left(\sup_{k\geqslant n} x_k\right)\,.$$

De même, si $m_n = \inf_{k\leqslant n} x_k$, on a

$$m_n \leqslant m_{n+1}\,, \quad n \geqslant 0\,,$$

et

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} m_n = \lim_{n\to\infty}\left(\inf_{k\geqslant n} x_k\right)\,.$$

On démontre que

$$-\infty \leqslant \liminf_{n\to\infty} x_n \leqslant \limsup_{n\to\infty} x_n \leqslant \infty$$

et que $\lim_{n\to\infty} x_n$ existe dans $\overline{\mathbb{R}}$ si et seulement si

$$\ell = \liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n\,;$$

on a alors $\ell = \lim_{n\to\infty} x_n$. Pour alléger la notation, nous écrivons

$$\limsup x_n\,, \quad \liminf x_n\,, \quad \lim x_n$$

en supprimant $n \to \infty$ ou en ajoutant à côté $n \to \infty$; ainsi, nous écrivons $x_n \to x$ $(n \to \infty)$ pour $\lim_{n\to\infty} x_n = x$. L'utilisation de lim sup, lim inf ne sera pas très fréquente dans ce volume ; néanmoins, leur maniement correct est essentiel pour tout étudiant en analyse.

Un ***intervalle ouvert*** de $\mathbb{R}$ est une partie de $\mathbb{R}$ (notée par $]a, b[$) déterminé par a, b dans $\overline{\mathbb{R}}$ avec $a < b$; on a

$$]a, b[= \{x \in \mathbb{R} : a < x < b\}\,;$$

l'intervalle ouvert $]a, b[$ s'appelle ***borné*** si a, b sont dans $\mathbb{R}$. Noter que $\mathbb{R} =]-\infty, \infty[$.

Avec $a < b$ dans $\overline{\mathbb{R}}$, on définit

$$[a, b[= \{x \in \mathbb{R} : a \leqslant x < b\} \quad (\text{si } a \in \mathbb{R})$$
$$]a, b] = \{x \in \mathbb{R} : a < x \leqslant b\} \quad (\text{si } b \in \mathbb{R})$$
$$[a, b] = \{x \in \mathbb{R} : a \leqslant x \leqslant b\} \quad (\text{si } a, b \in \mathbb{R}).$$

On dit que $[a, b]$ est un intervalle ***fermé***, ***borné*** dans $\mathbb{R}$ ou un intervalle ***compact*** dans $\mathbb{R}$; d'autres intervalles fermés de $\mathbb{R}$ sont

$$]-\infty, b],\ b \in \mathbb{R}\,; \quad [a, \infty[\,,\ a \in \mathbb{R}.$$

un ***intervalle de*** $\mathbb{R}$ est un ensemble d'un des quatres types $]a, b[$, $[a, b[$, $]a, b]$, $[a, b]$, définis ci-dessus. Noter que, selon notre définition, un intervalle contient toujours une infinité de points.

Soit X un ensemble. Si $f : X \to \mathbb{R}$, on dit que f est une ***fonction réelle*** sur X ; si $f : X \to \overline{\mathbb{R}}$, on dit que f est une ***fonction numérique*** sur X. Si $A \subset X$, on note par 1_A la ***fonction indicatrice de*** A ; c'est la fonction réelle 1_A sur X telle que

$$1_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \in A^c. \end{cases}$$

3 L'espace $\mathbb{R}^N$ $(N = 1, 2, \ldots)$

Soient $x = (x_1, \ldots, x_N)$, $y = (y_1, \ldots, y_N)$ deux éléments de $\mathbb{R}^N$ et $t \in \mathbb{R}$; on écrit

$$x + y = (x_1 + y_1, \ldots, x_N + y_N)\,, \quad tx = xt = (tx_1, \ldots, tx_N)$$
$$x \cdot y = \sum_{i=1}^{N} x_i y_i\,, \quad \|x\| = (x \cdot x)^{1/2}\,;$$

les nombres réels $x_1, \ldots, x_N$ s'appellent les ***coordonnées cartésiennes*** de x ; $x \cdot y$ s'appelle le ***produit scalaire*** (***euclidien***) entre x et y ; $\|x\|$ est la

norme euclidienne de x et $\|x-y\|$ est la ***distance euclidienne*** entre x et y. Si x, y de $\mathbb{R}^N$ sont différents de $0=(0,\ldots,0)$ l'***angle*** entre x et y est mesuré par l'unique nombre $\theta \in]-\pi\,,\,\pi]$ tel que

$$\cos\theta = \frac{x\cdot y}{\|x\|\,\|y\|}\,;$$

rappelons que l'existence et l'unicité de θ est garantie par l'***inégalité de Cauchy-Schwarz*** :

$$|x\cdot y| \leqslant \|x\|\,\|y\|\,,\quad x,\,y\in\mathbb{R}^N\,;$$

on dit que x est ***orthogonal*** à y ($x\perp y$) si $x\cdot y=0$; si $x\neq 0$, $y\neq 0$, alors $x\perp y$ si et seulement si l'angle θ entre x et y est égal à $\pi/2$.

$\mathbb{R}^N$ est un espace vectoriel réel muni de son produit scalaire euclidien. Les éléments de $\mathbb{R}^N$ seront tantôt appelés les ***vecteurs*** de $\mathbb{R}^N$ tantôt les ***points*** de $\mathbb{R}^N$. La ***base canonique*** de $\mathbb{R}^N$ est formée des vecteurs suivants :

$$e_1=(1,0,\ldots,0)\,,\ e_2=(0,1,0,\ldots,0)\,,\ \ldots\,,\ e_N=(0,0,\ldots,0,1)\,;$$

nous considérons $\{e_1,\ldots,e_N\}$ comme une ***base ordonnée*** de $\mathbb{R}^N$.

Soit $T:\mathbb{R}^N\to\mathbb{R}^M$ une application linéaire ; si $\{e_1,\ldots,e_N\}$ est la base canonique de $\mathbb{R}^N$ soit

$$Te_j=(t_{1j},\ldots,t_{Mj})\in\mathbb{R}^M\,,\quad j=1,\ldots,N\,.$$

La matrice $[t_{ij}]$ de ***type*** (M,N) (ce qui veut dire que la matrice a M lignes et N colonnes) possède les M lignes

$$(t_{i1},t_{i2},\ldots,t_{iN})\,,\quad i=1,\ldots,M\,,$$

et les N colonnes

$$(t_{1j},t_{2j},\ldots,t_{Mj})\,,\quad j=1,\ldots,N\,;$$

noter que les colonnes sont données par Te_j, $j=1,\ldots,N$. Nous dirons que $[t_{ij}]$ est la ***matrice associée*** à l'application linéaire $T:\mathbb{R}^N\to\mathbb{R}^M$; s'il n'y a aucune confusion à craindre, nous écrivons même

$$T=[t_{ij}]\,.$$

Rappelons que si $x=(x_1,\ldots,x_N)\in\mathbb{R}^N$, alors

$$Tx=\left(\sum_{j=1}^N t_{1j}x_j,\ldots,\sum_{j=1}^N t_{Mj}x_j\right)\in\mathbb{R}^M$$

ce qui correspond à la règle de multiplication « ligne par colonnes » entre la matrice $[t_{ij}]$ de type (M,N) et la matrice (vecteur) colonne

$$\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

de type $(N, 1)$ représentant x. Nous faisons la convention que

$$x = (x_1, \ldots, x_N) = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} ;$$

notons que l'on écrit Tx au lieu de $T(x)$ lorsque T est une application linéaire. Ainsi, avec $T = [t_{ij}]$, le vecteur Tx de $\mathbb{R}^M$ est simplement le produit matriciel

$$[t_{ij}] \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} .$$

Les matrices seront toujours notées par les crochets

$$[t_{ij}] , \quad \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} , \quad [x_1, \ldots, x_N] .$$

La matrice (vecteur) ligne de type $(1, N)$ sera notée comme

$$[x_1, \ldots, x_N] ,$$

les virgules étant utilisées pour séparer les termes ; si $T = [t_{ij}]$ est une matrice de type (M, N), ${}^tT$ représente la ***matrice transposée*** de T : on a

$${}^tT = [\tau_{ij}] \quad \text{avec } \tau_{ij} = t_{ji} , \; i = 1, \ldots, N , \; j = 1, \ldots, M$$

de sorte que ${}^tT$ est une matrice de type (N, M) ; de plus,

$$Tx \cdot y = x \cdot {}^tTy$$

où $x \in \mathbb{R}^N$, $y \in \mathbb{R}^M$, $T : \mathbb{R}^N \to \mathbb{R}^M$, ${}^tT : \mathbb{R}^M \to \mathbb{R}^N$. Selon nos notations, si $[x_1, \ldots, x_N]$ est une matrice ligne de type $(1, N)$,

$${}^t[x_1, \ldots, x_N] \equiv \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = (x_1, \ldots, x_N) = x$$

$${}^tx = [x_1, \ldots, x_N]$$

et si $x = (x_1, \ldots, x_N)$, $y = (y_1, \ldots, y_N)$ sont dans $\mathbb{R}^N$ alors

$$x \cdot y = {}^tx\, y = x{}^ty$$

où ${}^tx\, y$ est le produit matriciel de la matrice ligne ${}^tx$ par la matrice colonne y et de même pour $x{}^ty$. Rappelons que si

$$T : \mathbb{R}^N \longrightarrow \mathbb{R}^M , \quad S : \mathbb{R}^M \longrightarrow \mathbb{R}^P$$

sont deux applications linéaires avec

$$T = [t_{ij}], \quad S = [s_{ij}]$$

($[t_{ij}]$ matrice de type (M, N), $[s_{ij}]$ matrice de type (P, M)) alors l'application linéaire composée $S \circ T$, notée ST,

$$ST : \mathbb{R}^N \longrightarrow \mathbb{R}^P ,$$

est donnée par $ST = [u_{ij}]$ où

$$[u_{ij}] = [s_{ij}]\,[t_{ij}]$$

$$u_{ij} = \sum_{\alpha=1}^{M} s_{i\alpha} t_{\alpha j}, \quad i = 1, \ldots, P, \; j = 1, \ldots, N .$$

Ainsi ST correspond au produit matriciel $[s_{ij}]\,[t_{ij}]$.

Notons que si $T : \mathbb{R}^N \to \mathbb{R}^M$ est une application linéaire alors

$$\|Tx\| \leqslant C\|x\|, \quad x \in \mathbb{R}^N ,$$

où

$$C = \left(\sum_{i=1}^{M} \sum_{j=1}^{N} t_{ij}^2 \right)^{1/2}$$

si $T = [t_{ij}]$. En effet,

$$\|Tx\|^2 = \sum_{i=1}^{M} \left(\sum_{j=1}^{N} t_{ij} x_j \right)^2 \leqslant \sum_{i=1}^{M} \left\{ \sum_{j=1}^{N} t_{ij}^2 \right\} \left\{ \sum_{j=1}^{N} x_j^2 \right\} \leqslant C^2 \|x\|^2$$

d'après l'inégalité de Cauchy-Schwarz.

Rappelons que les applications linéaires

$$u : \mathbb{R}^N \longrightarrow \mathbb{R} = \mathbb{R}^1$$

s'appellent les ***formes*** ou ***fonctionnelles linéaires*** sur $\mathbb{R}^N$; à chaque u correspond un unique vecteur $(u_1, \ldots, u_N) \in \mathbb{R}^N$ tel que

$$ux = u \cdot x ;$$

aussi, la matrice de type (1,N) (matrice ligne) associée à u est $[u_1, \ldots, u_N]$; en écrivant $u = [u_1, \ldots, u_N]$ on aura

$$ux = [u_1, \ldots, u_N] \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = u \cdot x .$$

Soit $T : \mathbb{R}^N \to \mathbb{R}^M$, $T = [t_{ij}]$; on écrit aussi

$$T = [Te_1, \ldots, Te_N]$$

où $\{e_1, \ldots, e_N\}$ est la base canonique de $\mathbb{R}^N$ et $Te_1, \ldots, Te_N$ sont N vecteurs colonnes de $\mathbb{R}^M$.

Soit $\{\xi_1, \ldots, \xi_N\}$ une base ordonnée quelconque de $\mathbb{R}^N$ et $\{e_1, \ldots, e_N\}$ la base ordonnée canonique de $\mathbb{R}^N$. Il existe une unique transformation linéaire $T : \mathbb{R}^N \to \mathbb{R}^N$ telle que

$$Te_j = \xi_j\,, \quad j = 1, \ldots, N\,.$$

Si $T = [t_{ij}] = [\xi_1, \ldots, \xi_N]$ alors

$$\det T = \det[t_{ij}] \neq 0$$

(voir ci-après nos conventions concernant les déterminants des matrices carrées d'ordre N, c'est-à-dire de type (N, N)). Si $\det[t_{ij}] > 0$ la base ordonnée $\{\xi_1, \ldots, \xi_N\}$ est ***positive*** (ou ***positivement orientée***), sinon elle est ***négative*** (ou ***négativement orientée***). Noter que $[t_{ij}]$ est la matrice inversible qui permet de passer des coordonnées $y_1, \ldots, y_N$ de x dans la base $\{\xi_1, \ldots, \xi_N\}$ aux coordonnées cartésiennes $x_1, \ldots, x_N$ de x selon la formule

$$x_i = \sum_{j=1}^{N} t_{ij} y_j\,, \quad i = 1, \ldots, N\,;$$

en effet,

$$x = \sum_{i=1}^{N} x_i e_i = \sum_{i=1}^{N} y_i \xi_i = \sum_{i=1}^{N} y_i T e_i = T\left(\sum_{i=1}^{N} y_i e_i\right)$$

ce qui correspond à la relation matricielle

$$\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} t_{1\,1} & & t_{1N} \\ \vdots & & \vdots \\ t_{N1} & & t_{NN} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}.$$

Une matrice T de type (M, N) sera quelquefois notée comme

$$T = [\xi_1, \ldots, \xi_N]$$

où $\xi_1, \ldots, \xi_N$ sont les vecteurs colonnes de $\mathbb{R}^M$; ainsi, si $T = [t_{ij}]$ on aura

$$\xi_j = \begin{bmatrix} t_{1j} \\ \vdots \\ t_{Mj} \end{bmatrix} = (t_{1j}, \ldots, t_{Mj}) \in \mathbb{R}^M\,.$$

Lorsque $M = N$, on dit que T est une ***matrice carrée*** d'ordre N ; toutes les matrices telles que $T = [t_{ij}]$ dans ce volume seront ***réelles*** (c'est-à-dire les quantités t_{ij} sont réelles).

Un ***segment fermé*** d'extrémités x et y dans $\mathbb{R}^N$ (noté $[x, y]$) est l'ensemble

$$\{(1-t)x + ty : 0 \leqslant t \leqslant 1\} \subset \mathbb{R}^N ;$$

on définit un ***segment ouvert*** $]x, y[$ (ou semi-ouvert) dans $\mathbb{R}^N$ de manière évidente. Une partie E de $\mathbb{R}^N$ s'appelle ***convexe*** si pour tout x, y dans E, le segment fermé $[x, y]$ est contenu dans E.

Un ***pavé*** dans $\mathbb{R}^N$ est une partie A de $\mathbb{R}^N$ de la forme

$$A = I_1 \times \cdots \times I_N$$

où chaque I_j (les ***côtés*** de A) est un intervalle de $\mathbb{R}$; A est un ***pavé borné*** si les I_j sont bornés ; A est un ***pavé ouvert*** (resp. ***fermé***) si les I_j sont ouverts (resp. fermés) ; A est un ***pavé compact*** si les I_j sont des intervalles compacts de $\mathbb{R}$; A est un ***cube*** si A est un pavé borné et les longueurs des côtés I_j sont égales ; le sens des termes cubes ouverts, cubes fermés, est maintenant clair.

Soit $A \subset \mathbb{R}^N$, $A \neq \emptyset$; la quantité

$$\delta(A) = \sup\{\|x - y\| : x \in A,\ y \in A\}$$

s'appelle le ***diamètre*** de A ; si $\delta(A) < \infty$, on dit que A est une partie ***bornée*** de $\mathbb{R}^N$; $\emptyset$ est une partie bornée par convention ; une partie A de $\mathbb{R}^N$ est bornée si et seulement si A est contenue dans un cube (ou un pavé borné) ou encore que

$$\sup_{x \in A} \|x\| < \infty .$$

Soit $f : X \to \mathbb{R}^N$ où X est un ensemble quelconque ; alors

$$f(x) = \big(f_1(x), \ldots, f_N(x)\big), \quad x \in X ,$$

où $f_i : X \to \mathbb{R}$, $i = 1, \ldots, N$. Les fonctions réelles f_i sont les ***composantes*** de la ***fonction vectorielle*** f. Si $X \subset \mathbb{R}^N$ on dit que f est un ***champ vectoriel dans*** X. Si $N = 1$ on dit quelquefois que f est un ***champ scalaire*** dans X.

4 Permutations

Soit S_N l'ensemble de toutes les ***permutations*** de $\{1, \ldots, N\}$; un élément π de S_N est une bijection de $\{1, \ldots, N\}$ dans $\{1, \ldots, N\}$; π est une ***transposition*** entre i et j, $i \neq j$, si $\pi(i) = j$, $\pi(j) = i$ et $\pi(k) = k$ si $k \neq i, j$. *Toute permutation π est un produit d'un nombre fini de transpositions.* Voici une preuve par récurrence sur N : si $\pi \in S_{N+1}$ avec $\pi(N+1) = N+1$ alors π a la propriété indiquée par l'hypothèse de récurrence ; si $\pi(N+1) = k \neq N+1$ alors posons $\pi' = \tau\pi$ où τ est la transposition entre $N+1$ et k ; alors $\pi'(N+1) = N+1$ et on revient au cas précédent.

Si $\pi = \tau_1 \cdots \tau_r = \tau'_1 \cdots \tau'_s$, $\pi \in S_N$, *les* τ_i, τ'_i *étant des transpositions alors* $(-1)^r = (-1)^s$. En effet, posons

$$f(x_1, \ldots, x_N) = \prod_{1 \leqslant i < j \leqslant N} (x_j - x_i)$$

$$\pi f(x_1, \ldots, x_N) = f(x_{\pi(1)}, \ldots, x_{\pi(N)}) ;$$

on vérifie que

$$\pi f(x_1, \ldots, x_N) = (-1)^{\nu(\pi)} f(x_1, \ldots, x_N)$$

où

$$\nu(\pi) = \text{cardinalité de } \{(i,j) : 1 \leqslant i < j \leqslant N\,,\ \pi(i) > \pi(j)\}$$
$$= \text{nombre d'inversions produites par } \pi\,.$$

On conclut que

$$\pi_2(\pi_1 f) = (-1)^{\nu(\pi_1)+\nu(\pi_2)} f = (-1)^{\nu(\pi_2\pi_1)} f\,;$$

de plus, on vérifie que si τ est une transposition quelconque alors $\nu(\tau)$ est un nombre impair de sorte que $(-1)^{\nu(\tau)} = -1$; donc si

$$\pi = \tau_1 \cdots \tau_r = \tau'_1 \cdots \tau'_s$$

on a

$$(-1)^{\nu(\pi)} = (-1)^r = (-1)^s\,.$$

Pour $\pi \in S_N$ on pose $\varepsilon(\pi) = (-1)^{\nu(\pi)}$; $\varepsilon(\pi)$ s'appelle la ***signature*** de π ; si $\varepsilon(\pi) = 1$, π s'appelle une permutation ***paire*** et si $\varepsilon(\pi) = -1$, π s'appelle une permutation ***impaire***. Une permutation paire (resp. impaire) est le produit d'un nombre pair (resp. impair) de transpositions.

5 Déterminants

Associé à une matrice carrée d'ordre N, $T = [t_{ij}]$, est un nombre $\det T$ où $\det(T)$ (le ***déterminant*** de T) qui peut être défini comme suit :

$$\det(T) = \sum_{\pi} \varepsilon(\pi) t_{1\pi(1)} \cdots t_{N\pi(N)}$$

où la somme est étendue sur toutes les permutations π de $\{1, \ldots, N\}$ et $\varepsilon(\pi) = \pm 1$ est la signature de π. ·

Si S, T sont deux matrices carrées d'ordre N, on a

$$\det(ST) = \det(TS) = (\det S)(\det T)\,.$$

Une matrice carrée T d'ordre N est inversible si et seulement si $\det(T) \neq 0$; alors $T^{-1} = [s_{ij}]$ avec

$$s_{ij} = \frac{(-1)^{i+j} \det T_{ji}}{\det T}$$

où $T_{\alpha\beta}$ est la matrice carrée d'ordre $(N-1)$ obtenue de $T = [t_{ij}]$ en supprimant la ligne numéro α et la colonne numéro β ; $\det(T_{\alpha\beta})$ s'appelle un ***mineur d'ordre*** $(N-1)$ de T ou $\det(T)$.

On note

$$\det T = \begin{vmatrix} t_{11} & \dots & t_{1N} \\ \vdots & \ddots & \vdots \\ t_{N1} & & t_{NN} \end{vmatrix} \quad \text{si } T = [t_{ij}] .$$

Soit $T = [t_{ij}]$ une matrice de type (M, N) ; un ***mineur d'ordre*** k de T est un nombre

$$\det [t_{i_\alpha j_\beta}] , \quad \alpha , \beta = 1, \dots, k$$

où $1 \leqslant i_1 < \dots < i_k \leqslant N$, $1 \leqslant j_1 < \dots < j_k \leqslant N$; on a que $\operatorname{rang}(T) = r$ si tous les mineurs d'ordre $k > r$ de T sont nuls et il existe un mineur d'ordre r non nul. Rappelons que $T : \mathbb{R}^N \to \mathbb{R}^M$ est de rang r si et seulement si $\dim\big(T(\mathbb{R}^N)\big) = r$; on a $0 \leqslant r \leqslant \min(M, N)$. Aussi, $\operatorname{rang} T = \operatorname{rang}({}^tT)$ et $\operatorname{rang} T = r$ si et seulement si l'espace vectoriel engendré par les colonnes de T dans $\mathbb{R}^M$ (par les lignes de T dans $\mathbb{R}^N$) est de dimension r.

Si T est une matrice carrée d'ordre N,

$$\det(T) = \det({}^tT) .$$

Soit T une matrice carrée d'ordre N de la forme

$$T = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$$

où A est une matrice carrée d'ordre k, B une matrice carrée d'ordre $N-k$ et C est une matrice de type $(k, N-k)$; alors

$$\det T = (\det A)(\det B) .$$

Plus généralement, si

$$T = \begin{bmatrix} A_{11} & A_{12} & & A_{1r} \\ 0 & A_{22} & & A_{2r} \\ \vdots & & & \vdots \\ 0 & 0 & \dots & A_{rr} \end{bmatrix}$$

où A_{ii} est une matrice carrée d'ordre k_i alors

$$\det T = (\det A_{11}) \cdots (\det A_{rr}) .$$

Soit $T = [t_{ij}]$ une matrice carrée d'ordre N ; alors

$$\det T = \sum_{i=1}^{N} (-1)^{i+j} t_{ij} \det(T_{ij})$$

(développement de Laplace suivant la colonne numéro j)

$$\det T = \sum_{j=1}^{N} (-1)^{i+j} t_{ij} \det(T_{ij})$$

(développement de Laplace suivant la ligne numéro i). Plus généralement on a

$$\det T = \sum_{1 \leqslant i_1 < \ldots < i_k \leqslant N} (-1)^{i_1+\cdots+i_k+j_1+\cdots+j_k} \det T \begin{pmatrix} i_1 \ldots i_k \\ j_1 \ldots j_k \end{pmatrix} \det T' \begin{pmatrix} i_1 \ldots i_k \\ j_1 \ldots j_k \end{pmatrix}$$

pour $1 \leqslant j_1 < \ldots < j_k \leqslant N fixes$ où :

- $T\binom{i_j \ldots i_k}{j_1 \ldots j_k}$ est la matrice carrée d'ordre k obtenue de T en ne retenant que les lignes numéros $i_1 < \ldots < i_k$ et les colonnes numéros $j_1 < \ldots < j_k$,
- $T'\binom{i_j \ldots i_k}{j_1 \ldots j_k}$ est la matrice carrée d'ordre $(N-k)$ obtenue de T en supprimant les lignes numéros $i_1 < \ldots < i_k$ et les colonnes numéros $j_1 < \ldots < j_k$.

De même, on a pour $i_1 < \ldots < i_k$ fixés,

$$\det T = \sum_{1 \leqslant j_1 < \ldots < j_k \leqslant N} (-1)^{i_1+\cdots+i_k+j_1+\cdots+j_k} \det T \begin{pmatrix} i_1 \ldots i_k \\ j_1 \ldots j_k \end{pmatrix} \det T' \begin{pmatrix} i_1 \ldots i_k \\ j_1 \ldots j_k \end{pmatrix}.$$

Certaines des propriétés des déterminants sont développées directement dans le texte. Une référence utile est Gantmacher [21].

6 Espace $\mathbb{R}^3$

L'espace vectoriel $\mathbb{R}^3$ a la particularité de posséder un ***produit vectoriel*** qui est très commode pour certains calculs dans $\mathbb{R}^3$. Si $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$, on définit le produit vectoriel $x \times y$ entre x et y par le déterminant symbolique :

$$x \times y = \begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = (x_2y_3 - x_3y_2,\ x_3y_1 - x_1y_3,\ x_1y_2 - x_2y_1)\,.$$

Rappelons que $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$ et

$$e_1 \times e_2 = e_3\,, \quad e_2 \times e_3 = e_1\,, \quad e_3 \times e_1 = e_2\,.$$

Enumérons quelques propriétés du produit vectoriel :

- $x \times y$ est orthogonal à x et à y ;
- $\|x \times y\| = \|x\| \, \|y\| \sin\theta$ où θ est l'angle entre x et y, $x \neq 0$, $y \neq 0$: $\cos\theta = (x \cdot y)/\|x\| \, \|y\|$;
- $\|x \times y\|^2 + |x \cdot y|^2 = \|x\|^2 \|y\|^2$;
- $x \times y = -(y \times x)$, $x \mapsto x \times y$ est linéaire ;
- $a \times (b \times c) = b(a \cdot c) - c(a \cdot b)$;
- $(a \times b) \cdot c = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = [a, b, c]$ (produit mixte)

  où $a = (a_1, a_2, a_3)$, etc. ; la notation $[a, b, c]$ est standard et dans ce contexte elle ne crée aucune confusion avec la matrice carrée d'ordre 3 qui pourrait être définie par la même écriture ;
- $(a \times b) \times (c \times d) = b[c, d, a] - a[c, d, b] = c[a, b, d] - d[a, b, c]$;
- $(a \times b) \cdot (c \times d) = (a \cdot c)(b \cdot d) - (a \cdot d)(b \cdot c)$.

7 Cardinalité

Si X et Y sont deux ensembles, on dit que X et Y sont ***équipotents*** ou que X, Y sont de ***même cardinalité*** s'il existe une bijection de X sur Y. Si X est vide ou équipotent à $\{1, \ldots, n\}$ pour un $n \in \mathbb{N}^*$, on dit que X est un ensemble ***fini*** (de cardinalité 0 ou n). Si X n'est pas un ensemble fini, on dit que X est un ensemble ***infini*** ou que X et de ***cardinalité infinie***. Si X est équipotent à $\mathbb{N}$, on dit que X est ***dénombrable*** (ou que X est dénombrablement infini). Un ensemble X est ***au plus dénombrable*** s'il est dénombrable ou fini. $\mathbb{Z}$ et $\mathbb{Q}$ sont des ensembles dénombrables. On montre que $\mathbb{R}$ n'est pas dénombrable ; ainsi, on dit que $\mathbb{R}$ est ***non dénombrable*** (ou de cardinalité infinie non dénombrable). On peut démontrer les affirmations suivantes :

- soit $X = \bigcup_{i \in I} X_i$ avec $X_i \neq \emptyset$, $i \in I \neq \emptyset$; si I est dénombrable et les X_i sont au plus dénombrables alors X est dénombrable ; si I est fini et les X_i sont au plus dénombrables avec un au moins des X_i dénombrable alors X est dénombrable ;
- $\mathbb{R}$, $\mathbb{R}^N$ et tout sous-intervalle de $\mathbb{R}$ sont de même cardinalité.

8 Quelques fonctions réelles

Les fonctions trigonométriques (réelles) sont notées comme suit : sin, cos, tg, ctg, sec, cosec ; leurs fonctions réciproques sont notées comme suit : arc sin, arc cos, arc tg, etc. Rappelons que

$$\arcsin : [-1, 1] \longrightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

$$\arccos : [-1, 1] \longrightarrow [0, \pi]$$

$$\arctg : \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\longrightarrow \mathbb{R}$$

avec $\operatorname{arc\,tg}(\pm\pi/2) = \pm\infty$. La fonction exponentielle (réelle) $\exp : \mathbb{R} \to \mathbb{R}$ s'écrit aussi comme $\exp x = \mathrm{e}^x$. La fonction logarithmique (réelle) sera notée ln ; on a

$$\ln :]\,0\,, \infty[\longrightarrow \mathbb{R}$$

avec $\ln(\mathrm{e}^x) = x$, $x \in \mathbb{R}$, $\exp(\ln x) = x$, $x \in \mathbb{R}_+^*$. Les fonctions hyperboliques (réelles) sont notées comme suit : sh, ch, th, cth, sech, cosech avec

$$\operatorname{sh}(x) = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2}, \quad \operatorname{cosech}(x) = \frac{1}{\operatorname{sh}(x)} \text{ si } x \neq 0$$

$$\operatorname{ch}(x) = \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2}, \quad \operatorname{sech}(x) = \frac{1}{\operatorname{ch}(x)}$$

$$\operatorname{th}(x) = \frac{\operatorname{sh}(x)}{\operatorname{ch}(x)}, \quad \operatorname{cth}(x) = \frac{1}{\operatorname{th}(x)} \text{ si } x \neq 0\,.$$

9 Conventions diverses

Si $f : X \to Y$ nous écrivons quelquefois $f(x) \equiv c$ ou $f \equiv c$ pour indiquer que

$$f(x) = c\,, \quad x \in X\,,$$

ce qui veut dire que pour tout $x \in X$, $f(x)$ est égal à un élément fixe $c \in Y$.

Soient $f : X \to Y$, $g : X \to Y$; on exprimera ceci brièvement en écrivant $f, g : X \to Y$; une formule de type

$$f(x) = g(x)\,, \quad x \in A\,,$$

(où $A \subset X$) veut dire que pour tout x dans A on a $f(x) = g(x)$.

Soit X un ensemble ; si $a \in X$ et $b \in X$ on écrit brièvement : $a, b \in X$.

Soit $f : X \to \mathbb{R}$; si $f(x) \geqslant 0$, $x \in X$, on dit que f est (une fonction) ***positive*** ; si $f(x) > 0$, $x \in X$, on dit que f est ***strictement positive*** ; de même, pour f ***négative*** et ***strictement négative***.

Soit $f : I \to \mathbb{R}$, I un intervalle de $\mathbb{R}$; on dit que f est (une fonction) ***croissante*** (***décroissante***) si $x < y$ dans I entraîne que $f(x) \leqslant f(y)$ ($f(x) \geqslant f(y)$) ; f s'appelle ***strictement croissante*** (***strictement décroissante***) si $x < y$ dans I entraîne que $f(x) < f(y)$ ($f(x) > f(y)$).

Quelquefois (mais rarement) on a utilisé les symboles $\exists$ (il existe), $\forall$ (pour tout) et l'abréviation t.q. (tel que).

CHAPITRE 1

Topologie de $\mathbb{R}^N$

L'objectif principal de ce chapitre est d'introduire les notions topologiques de $\mathbb{R}^N$ qui sont essentielles pour une étude rigoureuse des fonctions de plusieurs variables réelles. Les fonctions d'une variable réelle étudiées dans les cours introductifs sont généralement définies sur des invervalles ; la nature géométrique d'un intervalle de la droite réelle est complètement déterminée par les considérations simples suivantes : l'intervalle est-il ouvert ou fermé ? s'il est fermé, est-il borné ou non ? Les fonctions de plusieurs variables réelles que l'on est amené à étudier naturellement sont, par contre, définies sur des parties de $\mathbb{R}^N$ plus variées, même si l'on veut se restreindre aux considérations les plus élémentaires comme celles de calcul de volume et d'aire des figures géométriques simples. Le fait est que pour les espaces $\mathbb{R}^N$, $N \geqslant 2$, il n'y a aucun analogue simple d'un intervalle de $\mathbb{R}$. On peut essayer de travailler avec des triangles (simplexes) ou des rectangles (pavés) « ouverts » ou « fermés » dans $\mathbb{R}^N$ mais l'on aperçoit que l'analogie préliminaire avec les intervalles de $\mathbb{R}$ disparaît rapidement. Par exemple, dans $\mathbb{R}$, la réunion d'intervalles non disjoints est toujours un intervalle, alors que la réunion de triangles ou de rectangles non disjoints dans $\mathbb{R}^2$ peut donner les figures les plus compliquées. En fait, une bonne généralisation de la réunion d'intervalles ouverts dans $\mathbb{R}$ est la notion d'un ensemble ouvert de $\mathbb{R}^N$. C'est là un concept de base pour toutes les considérations topologiques dans $\mathbb{R}^N$ (ou même, dans un espace quelconque) ; tous les autres concepts topologiques sont définis à partir de la notion d'ensembles ouverts. Parmi ceux-ci soulignons les suivants : adhérence, frontière, limite, compacité, connexité, continuité.

La plupart des considérations topologiques générales dont nous aurons besoin pour la suite du livre sont concentrées dans le présent chapitre. On conseille au débutant, pour une première lecture du livre, de bien comprendre la définition d'un ensemble ouvert dans $\mathbb{R}^N$ et celle d'une fonction continue. Bien entendu, il faudra, par la suite, revenir assez souvent à ce premier chapitre, pour approfondir ses connaissances de la topologie de $\mathbb{R}^N$.

1.1 PRÉLIMINAIRES

1.1.1 Définitions

Soit E un espace vectoriel (réel) et

$$p : E \to [0, \infty[$$

une application ayant les propriétés suivantes :

- positivité

$$p(x) = 0 \quad \text{équivaut à } x = 0 ; \tag{1.1}$$

- homogénéité

$$p(tx) = |t|p(x), \quad t \in \mathbb{R},\ x \in E ; \tag{1.2}$$

- inégalité du triangle

$$p(x + y) \leqslant p(x) + p(y), \quad x, y \in E. \tag{1.3}$$

Une telle application p s'appelle une ***norme*** dans E et E muni d'une norme p s'appelle un ***espace normé*** (E, p).

Pour $E = \mathbb{R}^N$, nous vérifions que

$$p(x) = \|x\| = (x \cdot x)^{1/2}, \quad x \in \mathbb{R}^N ,$$

et

$$p(x) = |x| = \max_i |x_i|, \quad x = (x_1, \ldots, x_N) \in \mathbb{R}^N$$

sont deux normes dans l'espace vectoriel $\mathbb{R}^N$. En effet, pour $p(x) = |x|$, les propriétés (1.1), (1.2), (1.3) sont immédiates ; pour $p(x) = \|x\|$, les propriétés (1.1) et (1.2) sont également immédiates alors que (1.3) est une conséquence du fait que $p(x)$ provient d'un produit scalaire dans $\mathbb{R}^N$ (inégalité de Cauchy-Schwarz, chap. 0).

La norme $\|x\|$ s'appelle la ***norme euclidienne*** dans $\mathbb{R}^N$; la norme $|x|$ s'appelle la ***norme maximum*** dans $\mathbb{R}^N$.

Si (E, p) est un espace normé, la formule

$$d(x, y) = p(x - y)$$

définit une application

$$d : E \times E \to [0, \infty[$$

qui satisfait les propriétés suivantes (pour tout x, y, z dans E) :

- positivité

$$d(x, y) = 0 \quad \text{équivaut à } x = y ; \tag{1.4}$$

• symétrie

$$d(x,y) = d(y,x)\ ; \tag{1.5}$$

• inégalité du triangle

$$d(x,y) \leqslant d(x,z) + d(z,y)\,. \tag{1.6}$$

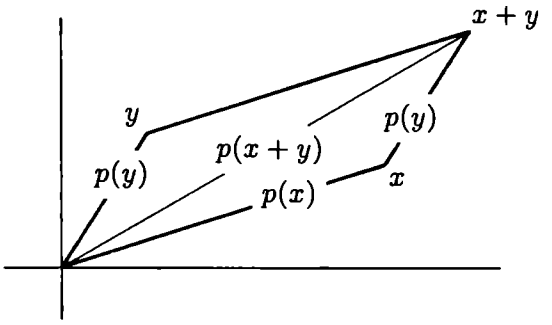


Fig. 1.1 Inégalité du triangle

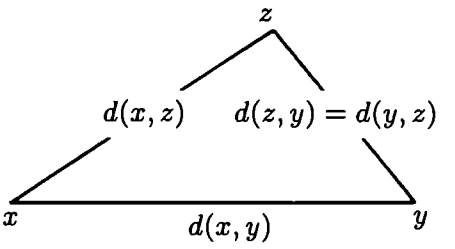


Fig. 1.2 Inégalité de distance

En effet, (1.4) est une conséquence de (1.1), (1.5) s'obtient de (1.2) avec $t = -1$ et (1.6) suit de (1.3) car

$$\begin{aligned} d(x,y) &= p(x-y) = p(x-z+z-y) \\ &\leqslant p(x-z) + p(z-y) = d(x,z) + d(z,y) \end{aligned}$$

Une application

$$d : E \times E \to [\,0\,,\,\infty\,[$$

ayant les propriétés (1.4), (1.5) et (1.6) s'appelle une ***distance*** (ou une ***métrique***) dans l'ensemble E (non nécessairement un espace vectoriel) et le couple (E,d) s'appelle un ***espace métrique*** ; $d(x,y)$ s'appelle la ***distance*** entre les points x, y de E.

Lorsque E est un espace vectoriel et

$$d(x,y) = p(x-y)\,, \quad x,\ y \in E\,,$$

où p est une norme dans E, on dit que la distance d est ***induite*** par la norme p. La distance

$$d(x,y) = \|x-y\|\,, \quad x,\ y \in \mathbb{R}^N$$

s'appelle la ***distance euclidienne*** (entre x, y) et elle est induite par la norme euclidienne. Par contre, nous ne donnerons pas un nom particulier à la distance

$$d(x,y) = |x-y|\,, \quad x,\ y \in \mathbb{R}^N$$

induite par la norme maximum.

Entre $|x|$ et $\|x\|$, on a les inégalités suivantes, pour $x \in \mathbb{R}^N$:

$$|x| \leqslant \|x\| \leqslant \sqrt{N}|x|\,. \tag{1.7}$$

La démonstration simple de (1.7) est laissée en exercice.

Une grande partie de ce que nous allons décrire dans la suite de ce chapitre est valable dans les espaces métriques généraux.

1.1.2 Exercices

1

(i) Dans un espace métrique quelconque (X, d), on a les inégalités suivantes :
(a) $\big|d(x,y) - d(y,z)\big| \leqslant d(x,z)$, x, y, $z \in X$;
(b) $d(x_1, x_n) \leqslant d(x_1, x_2) + \cdots + d(x_{n-1}, x_n)$
pour $x_1, \ldots, x_n$ dans X $(n \geqslant 3)$.

(ii) Dans un espace vectoriel normé (E, p), on a les inégalités suivantes :
(a) $\big|p(x) - p(y)\big| \leqslant p(x-y)$, x, y, $z \in E$;
(b) $p(x_1 + \cdots + x_n) \leqslant p(x_1) + \cdots + p(x_n)$, $x_1, \ldots, x_n \in E$.

2 Soit X un ensemble quelconque ; vérifier que, pour chaque $c > 0$,

$$d_c(x,y) = c \text{ si } x \neq y\,, \quad d_c(x,x) = 0$$

définit une distance dans X.

3 Pour $x = (x_1, \ldots, x_N)$ dans $\mathbb{R}^N$, posons

$$\|x\|_1 = \sum_{i=1}^{N} |x_i|\,.$$

Vérifier que $\|\cdot\|_1$ est une norme dans $\mathbb{R}^N$. On a les inégalités suivantes :

$$|x| \leqslant \|x\|_1 \leqslant N|x|\,, \quad \|x\| \leqslant \|x\|_1 \leqslant \sqrt{N}\|x\|\,.$$

4 Soit X un ensemble quelconque ; soit $B(X)$ l'ensemble de toutes les fonctions réelles, bornées dans X ; pour $f \in B(X)$, posons :

$$\|f\|_\infty = \sup_{x \in X} |f(x)|\,.$$

Vérifier que $\|\cdot\|_\infty$ est une norme dans l'espace vectoriel $B(X)$.

5 Soit $X = [0, 1]$; soit $C(X)$ l'ensemble de toutes les fonctions réelles, continues dans X ; pour f dans $C(X)$, posons

$$\|f\|_1 = \int_0^1 |f(x)|\,\mathrm{d}x\,.$$

Vérifier que $\|\cdot\|_1$ est une norme dans l'espace vectoriel $C(X)$; aussi,

$$\|f\|_1 \leqslant \|f\|_\infty\,, \quad f \in C(X)\,;$$

mais qu'une inégalité du type $(m > 0)$

$$\|f\|_1 \geqslant m\|f\|_\infty$$

ne peut pas être valable pour tout $f \in C(X)$.

1.2 ENSEMBLES OUVERTS

Soient $x \in \mathbb{R}^N$, $\varepsilon > 0$; le sous-ensemble de $\mathbb{R}^N$ défini par

$$B(x;\varepsilon) = \{y : \|y - x\| < \varepsilon\}$$

s'appelle la ***boule ouverte*** (***euclidienne***) de centre x et rayon ε. Une partie A de $\mathbb{R}^N$ s'appelle ***ouverte*** si pour tout x dans A, il existe un $\varepsilon > 0$ tel que

$$B(x;\varepsilon) \subset A\,.$$

Le sous-ensemble de $\mathbb{R}^N$ défini par

$$P(x;\varepsilon) = \{y : |y - x| < \varepsilon\}$$

s'appelle un ***cube ouvert*** de centre x et côtés de longueur 2ε.

Pour $N = 1$, $B(x;y)$ et $P(x;y)$ coïcident avec un ***intervalle ouvert*** centré à x et de longueur 2ε :

$$]\,x - \varepsilon\,,\, x + \varepsilon\,[\,, \quad x \in \mathbb{R}.$$

Pour $N = 2$, $B(x;\varepsilon)$ est un ***disque ouvert*** centré à x de rayon ε et $P(x;\varepsilon)$ est un ***carré ouvert*** centré à x dont les côtés sont de longueur 2ε : si $x = (x_1, x_2)$,

$$B(x;\varepsilon) = \{y = (y_1, y_2) : (y_1 - x_1)^2 + (y_2 - x_2)^2 < \varepsilon^2\}$$
$$P(x;\varepsilon) = \{y = (y_1, y_2) : |y_1 - x_1| < \varepsilon\,,\ |y_2 - x_2| < \varepsilon\}\,.$$

Pour $N = 3$, la terminologie adoptée pour $B(x;\varepsilon)$, $P(x;\varepsilon)$ correspond à l'usage de la géométrie élémentaire.

1.2.1 Proposition

(a) L'ensemble vide $\emptyset$ et l'espace entier $\mathbb{R}^N$ sont des ensembles ouverts.

(b) L'intersection d'un nombre fini de parties ouvertes est ouverte.

(c) La réunion d'une famille (finie ou infinie) de parties ouvertes est ouverte.

(d) Chaque boule ouverte est une partie ouverte.

(e) Un cube ouvert est ouvert ; une partie A de $\mathbb{R}^N$ est ouverte si et seulement si, pour tout x dans A, il existe un $\varepsilon > 0$ tel que

$$P(x;\varepsilon) \subset A\,.$$

DÉMONSTRATION

(a) Le fait que l'espace entier $\mathbb{R}^N$ est ouvert est évident ; le fait que l'ensemble vide est ouvert vient d'une interprétation logique de la définition d'une partie ouverte.

(b) Soient $A_1, \ldots, A_n$ les parties ouvertes, $A = \bigcap_i A_i$. Si $x \in A$, on a $x \in A_i$ pour tout i ; chaque A_i étant ouvert, il existe $\varepsilon_i > 0$ tel que :

$$B(x;\varepsilon_i) \subset A_i \ ;$$

si $\varepsilon = \min(\varepsilon_1, \ldots, \varepsilon_n)$ on aura

$$B(x;\varepsilon) \subset B(x;\varepsilon_i) \subset A_i \subset A\,,$$

ce qui démontre que A est une partie ouverte.

(c) Soit A_α, $\alpha \in J$, une famille quelconque de parties ouvertes, $A = \bigcup_\alpha A_\alpha$. Si $x \in A$, il existe un α tel que $x \in A_\alpha$; comme A_α est ouvert, il existe un $\varepsilon > 0$ tel que

$$B(x;\varepsilon) \subset A_\alpha \subset A$$

ce qui démontre que A est une partie ouverte.

(d) Soit $A = B(a;\rho)$; si $x \in A$, posons

$$\varepsilon = \rho - \|x - a\|\,.$$

On a $\varepsilon > 0$; si $y \in B(x;\varepsilon)$ on aura

$$\|y - a\| \leqslant \|y - x\| + \|x - a\| < \varepsilon + \|x - a\| = \rho\,,$$

ce qui démontre que

$$B(x;\varepsilon) \subset A\ ;$$

ceci établit que la boule ouverte A est une partie ouverte.

(e) Soit $A = P(a;\rho)$; si $x \in A$, posons

$$\varepsilon = \rho - |x - a|\,.$$

On a $\varepsilon > 0$; si $y \in P(x;\varepsilon)$ on aura

$$|y - a| \leqslant |y - x| + |x - a| < \varepsilon + |x - a| = \rho$$

ce qui démontre que

$$P(x;\varepsilon) \subset A\,.$$

De l'inégalité (1.7) on déduit que

$$B(x;\varepsilon) \subset P(x;\varepsilon) \subset A\,,$$

ce qui démontre que A est ouvert.

Pour la seconde partie de cette affirmation, nous observons d'abord que, de (1.7), on a les inclusions

$$B(x;\varepsilon) \subset P(x;\varepsilon) \subset B(x;\varepsilon\sqrt{N})\,.$$

Soit A une partie ouverte ; si $x \in A$, on a un $\varepsilon > 0$ tel que

$$B(x;\varepsilon) \subset A$$

d'où

$$P\left(x\,;\,\varepsilon/\sqrt{N}\right) \subset B(x;\varepsilon) \subset A\,.$$

D'autre part, si A est une partie ayant la propriété indiquée dans l'affirmation et $x \in A$, on a, pour un $\varepsilon > 0$,

$$B(x,\varepsilon) \subset P(x;\varepsilon) \subset A$$

ce qui montre que A est ouvert. ♦

REMARQUE. La définition d'une ***topologie*** dans un ensemble est la définition des *parties ouvertes* de celui-ci, satisfaisant les propriétés (a), (b), (c) de la proposition 1.2.1. On résume (e) de la proposition en disant que la topologie définie par la norme maximum $|\cdot|$ est identique avec la ***topologie euclidienne*** définie par la norme euclidienne $\|\cdot\|$; la démonstration montre que ceci provient de l'inégalité (1.7) ; celle-ci est décrite en disant que les deux normes, $|\cdot|$ et $\|\cdot\|$, sont ***équivalentes***.

1.2.2 Exercices

1 Un pavé ouvert de $\mathbb{R}^N$ est un ensemble ouvert (un pavé ouvert étant le produit cartésien d'intervalles ouverts). Chaque sous-ensemble ouvert est une réunion de pavés ouverts.

2 Soient $J_n = \,]-1/n\,,\,1/n\,[\,,\ n = 1, 2, \ldots$; chaque J_n est ouvert dans $\mathbb{R}$ mais $\bigcap_n J_n$ n'est pas ouvert. (Une intersection d'un nombre infini de parties ouvertes n'est pas nécessairement ouverte.)

3 L'ensemble $]\,0\,,\,1\,[\,\times\{0\}$ dans $\mathbb{R}^2$ n'est pas ouvert23051969.

4 Si A est ouvert dans $\mathbb{R}^M$ et B est ouvert dans $\mathbb{R}^N$, démontrer que $A \times B$ est ouvert dans $\mathbb{R}^{M\times N}$.

1.3 TERMINOLOGIE TOPOLOGIQUE

Un ***voisinage ouvert*** d'un point x est un ensemble ouvert contenant x.

Un point x est un point ***intérieur*** à un ensemble A, si A contient un voisinage ouvert de x ; l'ensemble formé de tous les points intérieurs à A s'appelle l'***intérieur*** de A et il sera noté $\mathring{A}$. Ainsi, x est dans $\mathring{A}$ si et seulement s'il existe une boule ouverte $B(x, \varepsilon)$ contenue dans A.

Un point x est ***adhérent*** à un ensemble A (ou x ***adhère à*** A) si tout voisinage ouvert V de x *rencontre* A, c'est-à-dire $V \cap A$ est non vide ; l'ensemble formé des points adhérents à A s'appelle l'***adhérence*** (aussi la ***fermeture***) de A et il sera noté $\overline{A}$.

Un point x est un ***point frontière*** d'un ensemble A, si x adhère à la fois à A et à A^c ; autrement dit, tout voisinage ouvert V d'un point frontière x de A, rencontre aussi bien A que A^c. L'ensemble formé des points frontières de A s'appelle la ***frontière*** de A et il sera noté $\operatorname{Fr} A$.

Un point x dans un ensemble A s'appelle un point ***isolé*** de A, s'il existe un voisinage ouvert V de x tel que $V \cap A$ est égal à $\{x\}$ (V ne rencontre A qu'en x).

Nous dirons que x est un ***point d'accumulation*** de l'ensemble A, si pour tout voisinage ouvert V de x, $V \cap A$ contient au moins un point de A autre que x ; il en suit immédiatement que $V \cap A$ contient une infinité de points de A. Un point x de $\mathring{A}$ est toujours un point d'accumulation de A ; si $x \in A$, alors x est un point d'accumulation de A, si et seulement si x n'est pas un point isolé de A ; si $x \in A^c$, alors x est un point d'accumulation de A, si et seulement si x adhère à A. On voit aussi que x est un point d'accumulation de A, si et seulement si x n'est pas un point isolé de l'ensemble $A \cup \{x\}$.

Un ensemble A s'appelle ***fermé*** si A^c est ouvert.

1.3.1 Quelques relations entre A, $\mathring{A}$, $\overline{A}$, $\operatorname{Fr} A$, $\operatorname{ext} A$

Les relations suivantes, pour un ensemble quelconque A, se déduisent facilement des définitions précédentes :

(i) $\mathring{A} \subset A \subset \overline{A}$;

(ii) $\operatorname{Fr} A = \operatorname{Fr} A^c = \overline{A} \cap \overline{A^c}$;

(iii) $\overline{A} = A \cup \operatorname{Fr} A = \mathring{A} \cup \operatorname{Fr} A$;

(iv) *les trois ensembles*

$$\mathring{A}, \quad \operatorname{Fr} A, \quad (A^c)^\circ$$

forment une partition de l'espace entier (ce qui veut dire qu'ils sont deux à deux disjoints et leur réunion est l'espace tout entier).

Pour voir ceci, il suffit d'observer que pour un point x quelconque, il n'y a que trois possibilités distinctes : ou bien il existe un voisinage ouvert V de x

contenu dans A (et, alors, $x \in \mathring{A}$), ou bien il existe un tel V contenu dans A^c (et, alors, $x \in (A^c)^\circ$), ou encore tout voisinage ouvert V de x rencontre A et A^c (et, alors, $x \in \operatorname{Fr} A$).

De (iii), on voit que

$$(\overline{A})^c = (A^c)^\circ . \tag{1.8}$$

L'ensemble défini par (1.8) s'appelle l'***extérieur*** de A ; il sera noté $\operatorname{ext} A$. Ainsi, l'intérieur, la frontière et l'extérieur d'un ensemble A, forment une partition de l'espace tout entier. De (1.8), on voit que $\operatorname{ext} A$ *est ouvert* et $\overline{A}$ *est fermé.*

(v) *Si* $A \subset B$, *alors*

$$\mathring{A} \subset \mathring{B}, \quad \overline{A} \subset \overline{B}.$$

1.3.2 Proposition

Pour tout ensemble A, $\mathring{A}$ *est un ensemble ouvert ; A est ouvert si et seulement si*

$$A = \mathring{A}.$$

L'ensemble $\mathring{A}$ *est la réunion de tous les ensembles ouverts contenus dans A. (On dit aussi que* $\mathring{A}$ *est le plus grand ensemble ouvert contenu dans A.)*

DÉMONSTRATION. Notons d'abord que, si A est ouvert, la définition même de $\mathring{A}$ donne que $A = \mathring{A}$; en particulier, si A est une boule ouverte, alors $A = \mathring{A}$.

Pour un ensemble quelconque A, si x est dans $\mathring{A}$, il existe une boule $B(x;\varepsilon)$ contenue dans A d'où (selon (v) du § 1.3.1) :

$$B(x;\varepsilon) \subset \mathring{A}$$

ce qui démontre que $\mathring{A}$ est ouverte. De ce qui précède, on conclut que A est ouvert exactement dans le cas où $A = \mathring{A}$.

Posons B pour la réunion de tous les ensembles ouverts contenus dans A ; évidemment, B est ouvert, étant une réunion de certains ensembles ouverts, d'où

$$\mathring{B} = B.$$

Comme $B \subset A$, on aura

$$\mathring{B} = B \subset \mathring{A}.$$

D'autre part, $\mathring{A} \subset B$ car $\mathring{A}$ est un ensemble ouvert contenu dans A, d'où on conclut que $B = \mathring{A}$. ♦

1.3.3 Proposition

(i) La famille d'ensembles fermés possède les propriétés suivantes :

(a) l'ensemble vide et l'espace entier sont fermés ;

(b) une intersection d'ensembles fermés est un ensemble fermé ;

(c) une réunion finie d'ensembles fermés est un ensemble fermé.

(ii) Pour tout ensemble A, $\overline{A}$ est fermé. Un ensemble A est fermé, si et seulement si $A = \overline{A}$; aussi, $\overline{A} = \overline{\overline{A}}$. L'ensemble $\overline{A}$ est l'intersection de tous les ensembles fermés contenant A. (On dit que $\overline{A}$ est le plus petit ensemble fermé contenant A.)

(iii) Pour tout ensemble A, $\mathrm{Fr}\, A$ est un ensemble fermé.

DÉMONSTRATION

(i) Un ensemble fermé étant le complémentaire d'un ensemble ouvert, ces propriétés se déduisent immédiatement de celles des ensembles ouverts (données dans la proposition 1.2.1) par passage aux complémentaires.

(ii) Le fait que $\overline{A}$ est fermé est une conséquence de (1.8) et de la proposition 1.3.2. Ainsi, si $A = \overline{A}$, A est fermé ; d'autre part, si A est fermé, alors A^c est ouvert et on aura, de la proposition 1.3.2 et (1.8),

$$A^c = (A^c)^\circ = (\overline{A})^c$$

ce qui établit, par complémentation, qu'alors, $A = \overline{A}$. La formule, $\overline{A} = \overline{\overline{A}}$, en est une conséquence.

Si B est l'intersection de tous les ensembles fermés contenant A, alors B est un ensemble fermé (cf. (i),(b)) qui contient A, d'où on a

$$\overline{B} = B \supset \overline{A}\,;$$

d'autre part, $\overline{A}$ est un ensemble fermé contenant A ce qui entraîne, selon la définition de B, que

$$B \subset \overline{A}\,;$$

au total, $B = \overline{A}$.

(iii) L'ensemble $\mathrm{Fr}\, A$ est l'intersection des ensembles fermés $\overline{A}$ et $\overline{A^c}$; donc $\mathrm{Fr}\, A$ est fermé. ♦

1.3.4 Remarque

Notons que toute la terminologie topologique est basée sur la notion d'ensembles ouverts. Ces derniers étaient définis dans la section 1.2 en utilisant les *boules ouvertes* obtenues de la norme euclidienne. Mais on a vu que nous aurions pu définir les ensembles ouverts en nous servant de la norme maximum (prop. 1.2.1(e)) ; géométriquement, ceci veut dire que nous pouvons remplacer les boules ouvertes de nos discussions par les cubes ouverts centrés de côtés égaux.

Si (X, d) est un espace métrique quelconque (sect. 1.1) on définit la boule ouverte $B(a; r)$, pour $a \in X$, $r > 0$, comme la partie de X donnée par

$$\{x \in X : d(x, a) < r\}\,.$$

A partir de ces boules ouvertes, on peut alors introduire toute la terminologie topologique exactement comme dans $\mathbb{R}^N$. Les propositions des sections 1.2 et 1.3 ainsi que leurs démonstrations restent valables dans ce cadre général.

1.3.5 Exercices

[1] Soit $A = \{1, 1/2, 1/3, \ldots\} \subset \mathbb{R}$. On a

$$\mathring{A} = \emptyset\,, \quad \overline{A} = A \cup \{0\}\,, \quad \operatorname{Fr} A = \overline{A}\,.$$

Tous les points de A sont les points isolés de A ; 0 est un point d'accumulation de A ; l'ensemble A est une réunion d'ensembles fermés sans être soi-même fermé.

[2] Soit A un sous-ensemble de cardinalité finie dans $\mathbb{R}^N$. Vérifier que

$$\mathring{A} = \emptyset\,, \quad \overline{A} = A\,, \quad \operatorname{Fr} A = A\,.$$

Tous les points de A sont les points isolés de A.

[3] Soit $A = \mathbb{R}^2 \setminus \{(x,0) : x \leqslant 0\}$. On a

$$\mathring{A} = A\,, \quad \overline{A} = \mathbb{R}^2\,, \quad \operatorname{Fr} A = \{(x,0) : x \leqslant 0\}\,.$$

Tous les points de $\mathbb{R}^2$ sont les points d'accumulation de A.

[4] Soit $A = B(a;r)$, $a \in \mathbb{R}^N$, $r > 0$. On a

$$\mathring{A} = A\,, \quad \overline{A} = \{x : \|x-a\| \leqslant r\}\,, \quad \operatorname{Fr} A = \{x : \|x-a\| = r\}\,.$$

(L'ensemble $\overline{A}$ s'appelle la ***boule fermée*** de centre a et de rayon r et sera notée $\overline{B}(a;r)$; l'ensemble $\operatorname{Fr} A$ s'appelle la ***sphère*** de centre a et rayon r et il sera noté $S(a;r)$; pour $N = 2$ on utilisera généralement les expressions ***disque ouvert*** pour décrire A, ***disque fermé*** pour $\overline{A}$ et ***cercle*** pour $\operatorname{Fr} A$.)

[5] Soit $A = B(a;r) \setminus \{a\}$, $a \in \mathbb{R}^N$, $r > 0$. On a

$$\mathring{A} = A\,, \quad \overline{A} = \overline{B}(a;r)\,, \quad \operatorname{Fr} A = S(a;r) \cup \{a\}\,.$$

Le point a est un point d'accumulation de A ; il est un point isolé de $\operatorname{Fr} A$. (*L'ensemble A de cet exercice sera noté $B'(a;r)$.*)

[6] Soient A, B deux sous-ensembles de $\mathbb{R}$: vérifier, pour l'ensemble $A \times B$ dans $\mathbb{R}^2$, les relations suivantes :

$$(A \times B)^\circ = \mathring{A} \times \mathring{B}\,, \quad \overline{(A \times B)} = \overline{A} \times \overline{B}$$
$$\operatorname{Fr}(A \times B) = (\operatorname{Fr} A \times \overline{B}) \cup (\overline{A} \times \operatorname{Fr} B)\,.$$

1.4 LIMITES

1.4.1 Limites des suites

Soit $\{x(n)\}$, $n \in \mathbb{N}$, une suite dans $\mathbb{R}^N$; on dit que la suite ***converge vers*** $x \in \mathbb{R}^N$ si

$$\lim_{n\to\infty} \|x(n) - x\| = 0 ;$$

on écrira : $x(n) \to x$ lorsque $n \to \infty$ ou $x(n) \to x$ $(n \to \infty)$ ou, simplement, $x(n) \to x$ ou encore

$$\lim_{n\to\infty} x(n) = x .$$

Si $x(n) \to x$, on dit que x est la ***limite*** de la suite ***convergente*** $\{x(n)\}$; cette terminologie (*la* limite) est justifiée par l'observation suivante : si $x(n) \to x$ et $x(n) \to y$ lorsque $n \to \infty$ alors $x = y$; en effet,

$$\|x - y\| \leqslant \|x - x(n)\| + \|x(n) - y\| \to 0 \quad (n \to \infty)$$

montre que $\|x - y\| = 0$ d'où $x = y$.

Si $x(n) \to x$ alors

$$\lim_{n\to\infty} \|x(n)\| = \|x\|$$

car

$$\big|\|x(n)\| - \|x\|\big| \leqslant \|x(n) - x\| \quad \text{(Ex. 1(ii), § 1.1.2).}$$

En utilisant la norme maximum et l'inégalité (1.7) on voit que $x(n) \to x$ si et seulement si

$$|x(n) - x| = \max_i |x_i(n) - x_i| \to 0$$

où $x(n) = \big(x_1(n), \ldots, x_N(n)\big)$, $x = (x_1, \ldots, x_N)$; on en déduit que $x(n) \to x$ si et seulement si

$$\lim_{n\to\infty} x_i(n) = x_i , \quad 1 \leqslant i \leqslant N .$$

On a aussi, alors, $|x(n)| \to |x|$ $(n \to \infty)$.

Une définition géométrique de la convergence de $x(n)$ vers x est la suivante : *pour tout voisinage ouvert V de x, il existe un n_0 tel que*

$$x(n) \in V \quad \text{pour } n \geqslant n_0 ;$$

il est évident que dans cette formulation, nous pouvons remplacer V par les voisinages ouverts spéciaux comme $B(x;\varepsilon)$ ou $P(x;\varepsilon)$.

La suite $\{x(n)\}$ est une ***suite de Cauchy*** si

$$\lim_{m,n\to\infty} \|x(m) - x(n)\| = 0\,,$$

ce qui veut dire que, pour tout $\varepsilon > 0$, il existe un n_0 tel que

$$\|x(m) - x(n)\| < \varepsilon \quad \text{si } m \geqslant n_0\,,\ n \geqslant n_0\ ;$$

nous écrivons cette propriété aussi comme suit :

$$x(m) - x(n) \to 0 \quad \text{lorsque } m\,,\ n \to \infty\,.$$

Une définition géométrique équivalente est la suivante : *pour tout voisinage ouvert V de $0 \in \mathbb{R}^N$, il existe un n_0 tel que*

$$x(m) - x(n) \in V \quad \text{si } m \geqslant n_0\,,\ n \geqslant n_0\ ;$$

de nouveau, nous pouvons remplacer V par des voisinages ouverts spéciaux $B(0;\varepsilon)$ ou $P(0;\varepsilon)$.

En passant à la norme maximum (1.7), on voit que $\{x(n)\}$ est une suite de Cauchy si et seulement si chaque suite réelle

$$\{x_i(n)\}_{n\in\mathbb{N}}\,, \quad i = 1, \ldots, N$$

est une suite de Cauchy où

$$x(n) = \big(x_1(n), \ldots, x_N(n)\big)\,.$$

Selon une propriété bien connue des nombres réels, chaque suite de Cauchy réelle est convergente ; on conclut, donc, qu'une suite de Cauchy $\{x(n)\}$ dans $\mathbb{R}^N$ converge vers $x = (x_1, \ldots, x_N)$ où

$$x_i = \lim_{n\to\infty} x_i(n)\,, \quad 1 \leqslant i \leqslant N\,.$$

Cette propriété de $\mathbb{R}^N$ que *toute suite de Cauchy dans $\mathbb{R}^N$ converge* est décrite en disant que $\mathbb{R}^N$ est un ***espace métrique complet***.

D'autre part, il est facile à établir que *toute suite convergente est une suite de Cauchy*.

1.4.2 Proposition

Soit $A \subset \mathbb{R}^N$. Un point x est dans $\overline{A}$ si et seulement si, pour au moins une suite $\{x(n)\}$ dans A,

$$\lim_{n\to\infty} x(n) = x\,.$$

L'ensemble A est fermé si et seulement si toute suite convergente, formée des points de A, converge vers un point de A.

DÉMONSTRATION. Il suffit d'observer que $x \in \overline{A}$ si et seulement si l'on peut choisir un $x(n)$ dans

$$B(x; 1/n) \cap A, \quad n = 1, 2, \ldots$$ ♦

1.4.3 Séries

Si $x(0), x(1), \ldots$ est une suite dans $\mathbb{R}^N$, on dit que la ***série***

$$\sum_{n=0}^{\infty} x(n) \quad \text{ou} \quad \sum_{n} x(n)$$

converge si la suite $\{S(n)\}$ des ***sommes partielles*** converge où

$$S(n) = x(0) + \cdots + x(n) ;$$

si S est limite de la suite $\{S(n)\}$, on dit que S est la ***somme*** de la série $\sum_n x(n)$ et on écrit

$$S = \sum_{n=0}^{\infty} x(n) \quad \text{ou} \quad S = \sum_{n \geqslant 0} x(n) \quad \text{ou} \quad S = \sum_{n} x(n).$$

On dit que la série $\sum_n x(n)$ ***converge absolument*** si

$$\sum_n \|x(n)\| < \infty$$

ce qui est équivalent à

$$\sum_n |x(n)| < \infty.$$

1.4.4 Proposition

(i) Une série $\sum_n x(n)$ absolument convergente est convergente et

$$\left\| \sum_n x(n) \right\| \leqslant \sum_n \|x(n)\|, \quad \left| \sum_n x(n) \right| \leqslant \sum_n |x(n)|.$$

(ii) Si $\sum_n x(n)$ converge alors $\lim_n x(n) = 0$.

DÉMONSTRATION

(i) Soit $S(n)$ la n-ième somme partielle de la série absolument convergente $\sum_n x(n)$; alors, si $n > m$,

$$\|S(n) - S(m)\| = \|x(m+1) + \cdots + x(n)\| \leqslant \|x(m+1)\| + \cdots + \|x(n)\|$$

d'où la convergence de la série $\sum_n \|x(n)\|$ permet de conclure que la suite $\{S(n)\}$ est une suite de Cauchy ce qui entraîne sa convergence ; les inégalités dans (i) sont des conséquences des inégalités du triangle pour les normes $\|\cdot\|$ et $|\cdot|$; par exemple,

$$\left\|\sum_{n=0}^{\infty} x(n)\right\| = \left\|\lim_{T\to\infty} \sum_{n=0}^{T} x(n)\right\| = \lim_{T\to\infty} \left\|\sum_{n=0}^{T} x(n)\right\|$$
$$\leqslant \lim_{T\to\infty} \sum_{n=0}^{T} \|x(n)\| = \sum_{n=0}^{\infty} \|x(n)\| .$$

(ii) De la relation

$$x(n) = S(n) - S(n-1)$$

on voit que, si $\sum_n x(n)$ converge vers S,

$$\|x(n)\| = \|S(n) - S + S - S(n-1)\|$$
$$\leqslant \|S(n) - S\| + \|S(n-1) - S\| \to 0 \quad (n \to \infty) .$$

1.4.5 Remarque

Notons de ce qui précède que l'étude d'une suite $\{x(n)\}$, $n \in \mathbb{N}$, ou d'une série $\sum_n x(n)$ dans $\mathbb{R}^N$ est équivalente à l'étude simultanée des N suites ou séries réelles :

$$\{x_i(n)\}_{n\in\mathbb{N}}, \quad \sum_n x_i(n), \quad i = 1, \ldots, N$$

où $x(n) = (x_1(n), \ldots, x_N(n))$. La convergence de la suite $\{x(n)\}$ équivaut à la convergence des N suites réelles $\{x_i(n)\}_{n\in\mathbb{N}}$ et

$$\lim_{n\to\infty} x(n) = (x_1, \ldots, x_N), \quad \text{où } x_i = \lim_{n\to\infty} x_i(n) .$$

Aussi $\sum_n x(n)$ converge (absolument) si et seulement si chacune des N séries réelles $\sum_n x_i(n)$, $i = 1, \ldots, n$, converge (absolument) et

$$\sum_n x(n) = \left(\sum_n x_1(n), \ldots, \sum_n x_N(n)\right) .$$

Ainsi, les propriétés des suites et séries réelles peuvent être étendues immédiatement aux suites et séries dans $\mathbb{R}^N$. Nous admettons, donc, certains résultats connus pour les suites et séries réelles, comme déjà établis pour les suites et séries dans $\mathbb{R}^N$.

La théorie de convergence des séries est équivalente à celle de convergence des suites ; ainsi une série $\sum_n x(n)$ converge si et seulement si la suite des

sommes partielles $\{S(n)\}$ converge, les relations entre les $x(n)$ et les $S(n)$ étant données par

$$S(n) = x(0) + \cdots + x(n)\,,\ n \geqslant 0\,; \quad x(n) = S(n) - S(n-1)\,,\ n \geqslant 1\,.$$

1.4.6 Limites des fonctions

Soit A un sous-ensemble de $\mathbb{R}^N$ et

$$f : A \to \mathbb{R}^M$$

une application de A dans $\mathbb{R}^M$; il s'agit alors de M fonctions réelles, $f_1, \ldots, f_M$ de N variables réelles $(x_1, \ldots, x_N) = x$ parcourant A :

$$f(x) = \big(f_1(x), \ldots, f_M(x)\big)\,, \quad x = \big(x_1, \ldots, x_N\big) \in A\,.$$

Soit a un point d'accumulation de A ; on dit que

$$\lim_{\substack{x \to a \\ x \in A}} f(x)$$

existe et est égal *à* $b \in \mathbb{R}^M$ *si pour tout* $\varepsilon > 0$, *il existe un* $\delta > 0$ *tel que*

$$\big\|f(x) - b\big\| < \varepsilon \quad \text{dès que } 0 < \|x - a\| < \delta\,,\ x \in A\,;$$

cette définition est équivalente à la suivante : *pour tout voisinage ouvert* V *de* b (dans $\mathbb{R}^M$), *il existe un voisinage ouvert* W *de* a (dans $\mathbb{R}^N$) *tel que*

$$f(x) \in V\,, \quad x \in (W \cap A) \setminus \{a\}\,.$$

(La vérification simple de cette équivalence est laissée comme un exercice impératif.) Notons que même si $a \in A$, ni l'existence de b ni sa valeur ne dépendent de $f(a)$. On démontre facilement que *l'existence de la limite entraîne son unicité* ; aussi, b *existe si et seulement si* $f(x_n) \to b$ *lorsque* $n \to \infty$, *pour toute suite* $\{x_n\}$ *dans* $A \setminus \{a\}$ *telle que* $x_n \to a$ $(n \to \infty)$.

Si a est un point isolé de A alors l'écriture

$$\lim_{\substack{x \to a \\ x \in A}} f(x)$$

signifie, par définition, $f(a)$. Ainsi, l'écriture

$$\lim_{\substack{x \to a \\ x \in A}} f(x)$$

sera utilisée seulement si a est un point d'accumulation de A (et alors a peut appartenir à A ou non) ou si a est un point isolé de A (donc $a \in A$).

Lorsque la limite

$$\lim_{\substack{x \to a \\ x \in A}} f(x) = b$$

existe, on écrira, pour simplifier la notation, comme suit : $f(x) \to b$ lorsque $x \to a$ dans A.

Le sous-ensemble A est dit ***non borné*** si

$$\sup\{\|x\| : x \in A\} = \infty .$$

Pour A non borné, on définit

$$\lim_{\substack{\|x\| \to \infty \\ x \in A}} f(x)$$

comme un élément $b \in \mathbb{R}^N$ tel que $\|f(x) - b\| \to 0$ lorsque $\|x\| \to \infty$, $x \in A$; autrement dit, pour tout $\varepsilon > 0$, il existe un $T > 0$ tel que

$$\|f(x) - b\| < \varepsilon , \quad x \in A , \ \|x\| > T ;$$

on écrira :

$$f(x) \to b \quad \text{lorsque } \|x\| \to \infty , \ x \in A .$$

Supposons que $M = 1$ ci-dessus, c'est-à-dire, f est une fonction réelle de $x \in A$; on écrit

$$\lim_{\substack{x \to a \\ x \in A}} f(x) = \infty \quad (\text{ou } +\infty)$$

(ou $f(x) \to \infty$ si $x \to a$, $x \in A$) si, pour $T > 0$, il existe un $\delta > 0$ tel que

$$f(x) > T , \quad 0 < \|x - a\| < \delta , \ x \in A .$$

Intuitivement, « $f(x) \to \infty$ si $x \to a$, $x \in A$, » veut dire simplement que les valeurs de f deviennent arbitrairement grandes lorsque x approche a, tout en restant dans $A \setminus \{a\}$. On définit et interprète « $f(x) \to -\infty$ si $x \to a$, $x \in A$, » de même manière.

Soit l'ensemble A un intervalle dans $\mathbb{R}$ et

$$f : A \to \mathbb{R}^M ;$$

si a est l'extrémité gauche de A, on définit $f(a_+)$ comme

$$\lim_{\substack{x \to a \\ x \in A \cap]a, \infty[}} f(x) ,$$

c'est-à-dire $f(x) \to f(a_+)$ si $x \to a$, $x \in A \setminus \{a\}$ (si la limite existe) ; de même, on définit $f(a_-)$ (si cela existe) lorsque a est l'extrémité droite de A.

Si A est une partie de $\mathbb{R}$ non bornée supérieurement, ce qui veut dire

$$\sup_{x \in A} x = \infty ,$$

et

$$f : A \to \mathbb{R}^M$$

on peut définir

$$\lim_{\substack{x \to \infty \\ x \in A}} f(x) = b\,;$$

si cette limite existe, on écrit $f(x) \to b$ si $x \to \infty$, $x \in A$. De même, si A est non bornée inférieurement, ce qui veut dire

$$\inf_{x \in A} x = -\infty\,,$$

on peut définir la notion de

$$\lim_{\substack{x \to -\infty \\ x \in A}} f(x) = b\,;$$

si cette limite existe, on écrit $f(x) \to b$ si $x \to -\infty$, $x \in A$.

D'autres limites peuvent être intéressantes à étudier aussi, mais on n'aura aucune difficulté à préciser les définitions appropriées à donner dans chaque cas. Il est vrai que l'on a des théories générales des limites qui contiennent tous les cas particuliers discutés ci-dessus. Mais nous n'aurons pas besoin de ces théories et les indications précédentes seront suffisantes pour la suite.

Les limites définies dans ce paragraphe ainsi que les limites des suites satisfont les propriétés usuelles des limites des suites réelles ; certaines de ces propriétés sont rappelées dans les exercices à la fin de cette section (§ 1.4.7).

1.4.7 Exercices

[1] Soient $\{x_n\}$, $\{y_n\}$, $n \in \mathbb{N}$, deux suites dans $\mathbb{R}^N$; si $x_n \to x$, $y_n \to y$ lorsque $n \to \infty$, démontrer que $z_n \to z$ lorsque $n \to \infty$ où

$$z_n = sx_n + ty_n\,, \quad z = sx + ty\,, \quad s,\, t \text{ réels}\,;$$

aussi, $x_n \cdot y_n \to x \cdot y$ lorsque $n \to \infty$.

[2] Soit $\{x_n\}$ une suite dans $\mathbb{R}^N$, $\{s_n\}$ une suite réelle, $n \in \mathbb{N}$; si $x_n \to x$ et $s_n \to s$ lorsque $n \to \infty$, démontrer que $z_n \to z$ lorsque $n \to \infty$ où

$$z_n = s_n x_n\,, \quad z = sx\,.$$

[3] Soit $A = \{(x, y) : 0 < y < 1/x\} \subset \mathbb{R}^2$; déterminer $\overline{A}$.

[4] Considérons une suite dans $\mathbb{R}^N$ de la forme

$$x_0\,,\ -x_0\,,\ x_1\,,\ -x_1\,,\ \ldots$$

Démontrer que la série $x_0 - x_0 + x_1 - x_1 + \cdots$ converge si et seulement si $x_n \to 0$ lorsque $n \to \infty$; elle converge absolument si et seulement si $\sum_n \|x_n\| < \infty$.

5 Soient f, g deux fonctions réelles dans $A \subset \mathbb{R}^N$ et a un point d'accumulation de A ; si $f(x) \to c$ et $g(x) \to d$ lorsque $x \to a$ dans A, démontrer que

$$f(x) + g(x) \to c + d\,, \quad f(x)g(x) \to cd\,, \quad \frac{f(x)}{g(x)} \to \frac{c}{d} \quad (d \neq 0)\,,$$

lorsque $x \to a$ dans A.

6 Soit

$$f(x,y) = \frac{xy}{x^2+y^2}\,, \quad (x,y) \neq (0,0)\,,\ (x,y) \in \mathbb{R}^2\,.$$

Voir que $f(1/n\,,\,1/n) = 1/2$, $n = 1, 2, \ldots$, mais que $f(x,0) = f(0,y) = 0$; conclure que, lorsque $(x,y) \to (0,0)$, $f(x,y)$ ne tend pas vers une limite.

7 Soit

$$f(x,y) = \frac{2x^3 - y^3}{x^2+y^2}\,, \quad (x,y) \neq (0,0)\,,\ (x,y) \in \mathbb{R}^2\,.$$

Démontrer que $f(x,y) \to 0$ lorsque $(x,y) \to (0,0)$. (Poser $x = \gamma\cos\theta$, $y = \gamma\sin\theta$ et étudier $\gamma \to 0$.)

8 Soient $f = (f_1, \ldots, f_M)$, $f : A \to \mathbb{R}^M$, a un point d'accumulation de A ; alors $f(x) \to b$, $b = (b_1, \ldots, b_M)$, lorsque $x \to a$ dans A si et seulement si $f_i(x) \to b_i$ lorsque $x \to a$ dans A, $i = 1, \ldots, M$.

9 Soit $\{x_n\}$, $n \in \mathbb{N}$, une suite dans $\mathbb{R}^N$ convergente vers $x \in \mathbb{R}^N$; toute sous-suite $\{x_{n(k)}\}$, $n(1) < n(2) < \ldots$, $n(k) \in \mathbb{N}$, est aussi convergente vers x.

10 Si A_i est un ensemble fermé dans $\mathbb{R}$, $i = 1, 2, \ldots, N$, alors $A_1 \times \cdots \times A_N$ est fermé dans $\mathbb{R}^N$. En particulier, un pavé fermé est fermé.

1.5 TOPOLOGIE INDUITE

1.5.1 Définitions

Soit X un sous-ensemble de $\mathbb{R}^N$. Si $A \subset X$, on dit que A est ***ouvert dans*** X ou A est ***ouvert relativement à*** X ou encore A est X-***ouvert*** si

$$A = V \cap X\,,$$

où V est un ensemble ouvert de $\mathbb{R}^N$; si A est ouvert dans X et $a \in A$ on dit que A est un ***voisinage ouvert de*** a ***relativement à*** X ou, plus brièvement, A est un ***voisinage*** X-***ouvert de*** a.

Un ensemble X-ouvert n'est pas nécessairement ouvert dans $\mathbb{R}^N$ (Ex. 2, § 1.5.3) ; si nous disons qu'un certain ensemble V (dans $\mathbb{R}^N$) est ***ouvert***, nous entendons que V ***est ouvert dans*** $\mathbb{R}^N$; en cas d'un doute quelconque, nous ajouterons la phrase « ouvert dans $\mathbb{R}^N$ » qui sera autrement omise.

Il est facile de vérifier que les ensembles X-ouverts satisfont les propriétés de stabilité indiquées dans (b) et (c) de la proposition 1.2.1 ; de plus, les ensembles $\emptyset$ et X sont toujours X-ouverts. On dit que les ensembles X-ouverts définissent la ***topologie induite*** de X. Toutes les notions topologiques définies dans $\mathbb{R}^N$ peuvent être *relativisées* à X en utilisant les ensembles X-ouverts.

A titre d'exemple, introduisons la notion d'adhérence relativement à X. Soit A un sous-ensemble de X ; un point x ***de*** X ***adhère à*** A ***relativement à*** X si tout voisinage X-ouvert de x rencontre A ; il est clair que ceci équivaut à la propriété suivante : pour tout voisinage ouvert V de x, x étant dans X,

$$A \cap V \neq \emptyset .$$

L'adhérence de A ***dans*** X (ou ***relativement à*** X) est l'ensemble $\overline{A}_X$ formé des points x de X qui adhère à A relativement à X ; on dira que $\overline{A}_X$ est la X-***adhérence*** de A.

De même, un sous-ensemble A de X s'appelle ***fermé dans*** X (***fermé relativement à*** X ou X-***fermé***) si $X \backslash A$ est X-ouvert. On voit immédiatement que A *est* X*-fermé si et seulement si*

$$A = F \cap X ,$$

où F *est une partie fermée* (dans $\mathbb{R}^N$).

Si A est X-fermé, on a même que

$$A = \overline{A} \cap X$$

car si

$$A = F \cap X , \quad F\text{-fermée}$$

alors

$$\overline{A} \subset F , \quad A = \overline{A} \cap A = \overline{A} \cap F \cap X = \overline{A} \cap X .$$

Ainsi, A *est* X*-fermé si et seulement si*

$$A = \overline{A} \cap X ;$$

ce qui veut dire que, *tout point de* X *qui adhère à* A *appartient* à A si et seulement si A est X-fermé.

En utilisant la caractérisation de $\overline{A}$ via les suites convergentes (prop. 1.4.2) on peut aussi énoncer que : A est X-fermé si et seulement si, *pour toute suite* $\{x_n\}$ *dans* A *convergeant vers un point* x *de* X, *on a que* $x \in A$ (étant entendu que $A \subset X$).

Plus généralement, on a la relation

$$\overline{A}_X = \overline{A} \cap X$$

valable pour toute partie A de X (Ex. 4, § 1.5.3).

On en déduit que A *est* X*-fermé si et seulement si*

$$\overline{A}_X = A$$

Il faut remarquer qu'un ensemble X-fermé n'est pas nécessairement fermé (dans $\mathbb{R}^N$) (Ex. 2, § 1.5.3).

Notons un critère simple, nécessaire et suffisant, pour qu'une partie A de $X \subset \mathbb{R}^N$ soit X-ouvert : A *est ouvert dans* X *si et seulement si, pour tout point* a *de* A*, il existe un voisinage* X*-ouvert* W *de* a *tel que*

$$W \subset A\,;$$

ceci revient à dire que A est X-ouvert, si et seulement si, pour tout point a de A, il existe un voisinage ouvert V de a tel que

$$V \cap X \subset A\,.$$

1.5.2 Remarques

Un sous-ensemble X de $\mathbb{R}^N$ est un espace métrique par rapport à la distance euclidienne restreinte à X. Comme indiqué à la fin de la section 1.1 et dans le paragraphe 1.3.4, la plupart des notions topologiques peuvent être définies dans tout espace métrique ; ces notions, traduites dans l'espace métrique X ($X \subset \mathbb{R}^N$, muni de la distance euclidienne), correspondent exactement aux notions topologiques relativisées introduites dans ce qui précède.

1.5.3 Exercices

[1] Si X est ouvert (fermé) dans $\mathbb{R}^N$ alors une partie A de X est ouverte (fermée) dans X si et seulement si A est ouvert (fermé) dans $\mathbb{R}^N$.

[2] Soit $X = [0\,,\,1\,[\subset \mathbb{R}$; l'intervalle $[0\,,\,1/2\,[$ est ouvert dans X sans être ouvert dans $\mathbb{R}$; l'intervalle $[1/2\,,\,1\,[$ est fermé dans X sans être fermé dans $\mathbb{R}$.

[3] Soit $X \subset \mathbb{R}^N$; vérifier que la famille d'ensembles fermés relatifs à X satisfait la proposition 1.3.3(i).

[4] Vérifier la formule $\overline{A}_X = \overline{A} \cap X$, $A \subset X \subset \mathbb{R}^N$.

1.6 FONCTIONS CONTINUES

Les fonctions continues sont des objets centraux dans toutes les théories topologiques. Du point de vue pratique, une fonction continue

$$f : A \to \mathbb{R}^M , \quad A \subset \mathbb{R}^N , \; f = (f_1, \ldots, f_M)$$

représente une modélisation de M quantités physiques $f_1, \ldots, f_M$ qui dépendent de N paramètres $(x_1, \ldots, x_N)$ dans A, de manière continue, c'est-à-dire si x, y sont dans A et sont très près l'un de l'autre (ce qui veut dire que $\|x-y\|$ dans $\mathbb{R}^N$ est petit) alors les valeurs $f(x)$, $f(y)$ sont aussi très proches l'une de l'autre (dans le sens que $\|f(x)-f(y)\|$ dans $\mathbb{R}^M$ est petit) ; ceci équivaut à la continuité des chacunes des fonctions réelles f_i individuellement.

1.6.1 Définitions

Soit $A \subset \mathbb{R}^N$ et $f : A \to \mathbb{R}^M$. On dit que f est ***continue*** en un point $a \in A$, si pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que

$$\|f(x) - f(a)\| < \varepsilon , \quad x \in A , \; \|x - a\| < \delta ;$$

autrement dit,

$$f\{B(a;\delta) \cap A\} \subset B\big(f(a);\varepsilon\big) .$$

Il est clair que cette définition est équivalente à la suivante : f est continue en un point $a \in A$, si pour tout voisinage ouvert V de $f(a)$ (dans $\mathbb{R}^M$) il existe W, un voisinage A-ouvert de a, tel que

$$f(W) \subset V .$$

On dit que f est ***continue dans*** A et on écrit

$$f \in C(A;\mathbb{R}^M)$$

si f est continue en chaque point de A ; si $M = 1$, on écrira $C(A)$ pour $C(A;\mathbb{R}^M)$.

Soient $A \subset \mathbb{R}^N$, $B \subset \mathbb{R}^M$, $f : A \to B$. On dit que f est un ***homéomorphisme*** entre A et B si f est *bijective*, *continue*, et $f^{-1} : B \to A$ est *continue*. Les ensembles A et B sont dits ***homéomorphes*** s'il existe un homéomorphisme entre A et B.

Soit $A \subset \mathbb{R}^N$, $f : A \to \mathbb{R}^M$. On dit que f est ***uniformément continue dans*** A si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que, pour tout point $a \in A$,

$$\|f(x) - f(a)\| < \varepsilon , \quad x \in A , \; \|x - a\| < \delta ;$$

ce qu'il faut observer dans cette définition c'est que le choix de δ ne dépend pas du point $a \in A$. Il est évident qu'une fonction uniformément continue dans A est continue dans A mais la réciproque est en défaut, en général (voir Ex. 9, § 1.6.4 ; prop. 1.7.3(iii)).

1.6.2 Proposition

Soit $A \subset \mathbb{R}^N$, $f : A \to \mathbb{R}^M$, $f = (f_1, \ldots, f_M)$.

(i) f *est continue au point* a *si et seulement si*

$$\lim_{\substack{x \to a \\ x \in A}} f(x) = f(a)$$

ou encore, si pour toute suite $\{x_n\}$ *dans* A *telle que* $x_n \to a$ *alors* $f(x_n) \to f(a)$ *lorsque* $n \to \infty$.

La continuité de f *au point* a *équivaut à la continuité de chacune des* M *fonctions réelles* $f_1, \ldots, f_M$ *au point* a. *En particulier,* $f \in C(A; \mathbb{R}^M)$ *si et seulement si* $f_i \in C(A)$, $i = 1, \ldots, M$.

(ii) f *est continue dans* A *si et seulement si l'image inverse par* f *de tout ensemble ouvert (fermé) dans* $\mathbb{R}^M$ *ou dans* $B = f(A)$ *est ouverte (fermée) dans* A.

(iii) Soit $g : B \to \mathbb{R}^P$ *tel que* $B \supset f(A)$. *Si* f *est continue au point* $a \in A$ *et* g *est continue au point* $b = f(a)$ *dans* B *alors*

$$g \circ f : A \to \mathbb{R}^P$$

est continue au point $a \in A$. *En particulier, si*

$$f \in C(A, \mathbb{R}^M), \quad g \in C(B, \mathbb{R}^P), \quad B \supset f(A)$$

alors $g \circ f \in C(A, \mathbb{R}^P)$.

DÉMONSTRATION

(i) C'est une conséquence immédiate des définitions (§ 1.6.1 ; Ex. 8, § 1.4.7).

(ii) Soit f continue dans A et V une partie ouverte de $\mathbb{R}^M$; posons

$$W = f^{-1}(V) = \{x \in A : f(x) \in V\}.$$

Si $x \in W$, $y = f(x)$ est dans V ; comme V est un voisinage ouvert de y et f est continue au point $x \in A$, il existe un voisinage A-ouvert W_x de x tel que

$$f(W_x) \subset V ;$$

d'où

$$W_x \subset f^{-1}(V) = W ,$$

ce qui montre que W est ouvert dans A.

Supposons maintenant que $f^{-1}(V)$ est ouvert dans A, pour tout choix d'une partie ouverte V dans $\mathbb{R}^M$; soit $a \in A$ et $f(a) = b$; si V est un voisinage ouvert de b, $W = f^{-1}(V)$ est un ouvert dans A, donc un voisinage A-ouvert de a tel que

$$f(W) \subset V\,,$$

d'où la continuité de f au point a.

L'affirmation concernant les ensembles fermés se déduisent par complémentation ; en effet, si F est fermé dans $\mathbb{R}^M$, F^c est ouvert ; si f est continue dans A, $f^{-1}(F^c)$ sera ouvert dans A ; donc l'ensemble

$$A \setminus f^{-1}(F^c)$$

sera fermé dans A ; comme

$$f^{-1}(F) = A \setminus f^{-1}(F^c)\,,$$

on conclut que $f^{-1}(F)$ est fermé dans A. Le même raisonnement établit la réciproque qui dit que si $f^{-1}(F)$ est fermé dans A, pour tout choix de l'ensemble fermé F dans $\mathbb{R}^M$ alors f est continue dans A. Les raisonnements nécessaires pour traiter les cas de $f^{-1}(E)$, où E est ouvert (fermé) dans $B = f(A)$, sont exactement analogues. En effet, si E est ouvert dans B, E est de la forme $V \cap B$, V étant ouvert dans $\mathbb{R}^M$ et si E est fermé dans B alors E est de la forme $F \cap B$, F étant fermé dans $\mathbb{R}^M$; les raisonnements précédents s'appliquent alors sur V et F.

(iii) Posons $b = f(a)$, $c = g(b)$; soit U un voisinage ouvert de c dans $\mathbb{R}^P$; par la continuité de g au point b, on obtient un voisinage B-ouvert V' de b tel que

$$g(V') \subset U$$

soit V un ensemble ouvert de $\mathbb{R}^M$ tel que

$$V' = V \cap B\,;$$

la continuité de f au point a garantit l'existence d'un voisinage A-ouvert W de a tel que

$$f(W) \subset V\,,$$

ce que veut dire que

$$f(W) \subset V' = V \cap B$$

car $f(W) \subset B$; on a, ainsi,

$$g \circ f(W) \subset g(V') \subset U$$

ce qui établit la continuité de $g \circ f$ au point a. ♦

1.6.3 Remarques

On peut résumer la proposition 1.6.2, en termes généraux, comme suit : la composition des applications continues donne des applications continues ; les applications continues sont caractérisées par la propriété que les images inverses des ensembles ouverts (fermés) sont ouverts (fermés).

Notons que l'image directe d'un ensemble ouvert (fermé) sous une application continue n'est pas nécessairement ouvert (fermé) ; une application est dite ***ouverte*** (***fermée***) si l'image directe de chaque ensemble ouvert (fermé) est ouvert (fermé) (Ex. 5 et 7, § 1.6.4).

Les applications coordonnées

$$\pi_1(x_1, \ldots, x_N) = x_i\,, \quad i = 1, \ldots, N\,,$$

sont évidemment continues ; on le voit directement des définitions ou via (i) de la proposition 1.6.2. Si f, g sont réelles et continues dans un ensemble $A \subset \mathbb{R}^M$ alors $f+g$, $f \cdot g$ sont continues dans A ; aussi, f/g est continue dans A pourvu que $g(x) \neq 0$, $x \in A$. On peut démontrer ces affirmations via la partie (i) de la proposition 1.6.2 et les propriétés usuelles des limites ; on peut aussi les démontrer comme suit : l'application

$$x \mapsto \big(f(x), g(x)\big) \in \mathbb{R}^2\,, \quad x \in A\,,$$

est continue ; si

$$\theta : \mathbb{R}^2 \to \mathbb{R}$$

est continue, l'application composée

$$x \mapsto \theta\big(f(x), g(x)\big)$$

sera continue ; en posant

$$\theta(x, y) = x + y \text{ ou } xy$$

on obtient la continuité de $f+g$ ou fg via la partie (iii) de la proposition 1.6.2. En considérant

$$\theta(x, y) = \frac{x}{y}\,, \quad y \neq 0$$

on obtient, de même, la continuité de f/g.

De ce qui précède, on établit la continuité des fonctions ***polynomiales*** dans $\mathbb{R}^N$; ces fonctions sont des combinaisons linéaires des ***monômes*** de la forme

$$x_1^{n_1}\, x_2^{n_2}\, \cdots\, x_N^{n_N}\,, \quad n_1, \ldots, n_N \text{ entiers} \geqslant 0\,;$$

le ***degré*** de ce monôme est $n_1 + \cdots + n_N$; le degré d'un polynôme est, par définition, le degré le plus grand des monômes qui apparaissent dans la présentation du polynôme comme une combinaison linéaire, de coefficients non nuls, des monômes.

Une fonction ***rationnelle*** dans $\mathbb{R}^N$ est une fonction f de type P/Q où P, Q sont des fonctions polynomiales, définie dans l'ensemble D ou Q est non nulle ; f est alors continue sur D.

Deux avertissements :

(i) si f est une fonction définie dans un ensemble X et $A \subset X$ alors il se peut que la restriction de f à A soit continue sans que f soit continue aux points de A en tant que fonction définie dans X (Ex. 8, § 1.6.4) ;

(ii) une fonction de deux variables peut très bien être continue séparément en chaque variables sans qu'elle soit continue en tant que fonction des deux variables dans leur ensemble (Ex. 10, § 1.6.4).

1.6.4 Exercices

1 Considérer les expressions suivantes :

(i) $f(x,y) = x/y$;
(ii) $f(x,y) = (y, \sqrt{xy})$;
(iii) $f(x,y,z) = (x+y+z,\ x^2y,\ y^2/z)$;
(iv) $f(x,y) = (x,\ x^2y,\ x^3y^2,\ x^4y^3)$.

Placer chacune de ces expressions dans un cadre $f : A \to \mathbb{R}^M$, $A \subset \mathbb{R}^N$ et discuter leur continuité.

2 Soit $f : A \to \mathbb{R}^M$, $A \subset \mathbb{R}^N$, $a \in A$; démontrer que f est continue au point a si et seulement si $f(x_n) \to f(a)$ dès que $x_n \in A$, $x_n \to a$ $(n \to \infty)$. Si a est un point isolé, alors f est continue au point a.

3 Soit $f(x) = a \cdot x$ $(a \in \mathbb{R}^N)$; $f : \mathbb{R}^N \to \mathbb{R}$ est continue. En déduire que les demi-espaces ouverts (fermés)

$$\{x : a \cdot x > \alpha\} \quad (\{x : a \cdot x \geqslant \alpha\})$$

sont des ensembles ouverts (fermés) et que l'hyperplan

$$\{x : a \cdot x = \alpha\}$$

est un ensemble fermé dans $\mathbb{R}^N$.

4 Soit $f : A \to \mathbb{R}^M$, $A \subset \mathbb{R}^N$, $a \in A$; si f est continue au point a et $f(a) \neq 0$, démontrer que, pour un $\varepsilon > 0$ approprié, $f(x) \neq 0$ pour x dans $B(a;\varepsilon) \cap A$. En déduire que, si $g : A \to \mathbb{R}^M$ et f, g sont continues au point a avec $f(a) \neq g(a)$ alors $f(x) \neq g(x)$ pour x dans un voisinage A-ouvert de a.

5 L'application coordonnée

$$\theta(x,y) = x\,, \quad (x,y) \in \mathbb{R}^2\,,$$

est une application continue ouverte mais non fermée (considérer $\theta(F)$ où $F = \{(x,y) : xy = 1\}$). L'application $\widetilde{\theta}(x,y) = (x,0)$ de $\mathbb{R}^2$ dans $\mathbb{R}^2$ est continue mais ni ouverte, ni fermée.

6 Soit $f : A \to B$, $A \subset \mathbb{R}^N$, $B \subset \mathbb{R}^M$; si f est une bijection telle que, pour tout A-ouvert (A-fermé) V, $f(V)$ est un B-ouvert (B-fermé) et, pour tout B-ouvert (B-fermé) W, $f^{-1}(W)$ est un A-ouvert (A-fermé) alors f est un homéomorphisme.

7 Soit $f(t) = (\cos t, \sin t)$, $t \in A = [\,0\,,\,2\pi\,[$; l'application

$$f : A \to \mathbb{R}^2$$

est injective et continue mais f n'est pas un homéomorphisme entre A et $f(A)$ car f^{-1} n'est pas continue.

8 Soit $A \subset \mathbb{R}^N$ et $f = 1_A$; f est continue en un point x si et seulement si $x \in \mathring{A}$ ou $x \in \operatorname{ext} A$; si A est tel que $\overline{A} = \mathbb{R}^N$ et $\overline{A^c} = \mathbb{R}^N$ (par exemple, si A est l'ensemble des $x = (x_1, \ldots, x_N)$ avec tous les x_i rationnels) alors f est discontinue partout ; par contre, si A est non vide, la fonction $\widetilde{f} = f|_A$ (la restriction de f à A) est toujours continue (car constamment égale à 1).

9 Soit $f : \mathbb{R} \to \mathbb{R}$; si

$$\big|f(x) - f(y)\big| \leqslant \varphi\big(|x-y|\big)\,, \quad x,\, y \in \mathbb{R},$$

où

$$\varphi : [\,0\,,\,\infty\,[\to \mathbb{R}, \quad \lim_{t\to 0} \varphi(t) = \varphi(0) = 0\,,$$

alors f est uniformément continue. En conclure, que

$$f(x) = cx\,, \quad f(x) = |x|\,, \quad f(x) = \sin x$$

donnent des fonctions uniformément continues. Par contre, $f(x) = x^2$ donne une fonction qui n'est pas uniformément continue.

10 Soit

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{si } (x,y) \in \mathbb{R}^2,\ (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0)\,. \end{cases}$$

Pour chaque y fixe, la fonction $x \mapsto f(x,y)$ est continue dans $\mathbb{R}$; pour chaque x fixe, la fonction $y \mapsto f(x,y)$ est continue dans $\mathbb{R}$; mais f, en tant que fonction de (x,y), n'est pas continue au point $(0,0)$ (Ex. 6, § 1.4.7).

11 Une application linéaire $u : \mathbb{R}^N \to \mathbb{R}^M$ quelconque est continue.

1.7 COMPACITÉ

Les ensembles *fermés* et *bornés* dans $\mathbb{R}^N$ possèdent plusieurs propriétés agréables ; du point de vue moderne, on les caractérisent en passant par une propriété de recouvrement que nous formulons maintenant.

Soit A une partie de $\mathbb{R}^N$. On dit que la famille d'ensembles $\{V_\alpha\}$, $\alpha \in J$, est un ***recouvrement*** de A si

$$A \subset \bigcup_\alpha V_\alpha \ ;$$

il s'appelle un ***recouvrement ouvert*** de A si, de plus, chaque V_α est ouvert. On dit que A est ***compact*** si, de tout recouvrement ouvert $\{V_\alpha\}$ de A, on peut extraire un recouvrement fini $V_{\alpha_1}, \ldots, V_{\alpha_n}$. On dit que A est ***borné*** si, pour un $R > 0$,

$$A \subset B(0; R) \ ;$$

ceci revient à dire que

$$\sup\{\|x\| : x \in A\} < \infty \, .$$

Remarquons que A est borné, si et seulement si

$$\delta(A) = \sup\{\|x - y\| : x, y \in A\} < \infty \ ;$$

$\delta(A)$ s'appelle le ***diamètre*** de l'ensemble A.

1.7.1 Proposition

Les affirmations suivantes concernant une partie A de $\mathbb{R}^N$ sont équivalentes :

(a) (propriété de Borel-Heine-Lebesgue) A est compact ;

(b) (principe de Bolzano-Weierstrass) de toute suite dans A, on peut extraire une sous-suite qui converge vers un point de A ;

(c) A est fermé et pour tout $\varepsilon > 0$, il existe une partie finie S_ε de A telle que tout point de A est à une distance inférieure à ε d'un point de S_ε ; autrement dit,

$$A \subset \bigcup \{B(s; \varepsilon) : s \in S_\varepsilon\} \ ;$$

(d) A est fermé et borné.

DÉMONSTRATION

(a)⇒(b) Démontrons d'abord que A est fermé, c'est-à-dire A^c est ouvert. Si $x \in A^c$ et $a \in A$, il existe un voisinage ouvert V_a de a tel qu $\overline{V}_a$ ne contient pas x ; il suffit de prendre V_a comme $B(a;r)$ avec

$$0 < r < \|x - a\| .$$

Puisque les V_a, $a \in A$, forment un recouvrement ouvert de A, on peut en extraire, selon (a), un recouvrement fini

$$V_{a_1}, \ldots, V_{a_n} ;$$

si F est la réunion des $\overline{V}_{a_j}$, $j = 1, \ldots, n$, alors F est fermé et

$$F \supset A , \quad x \notin F ,$$

ce qui veut dire que l'ensemble ouvert $F^c = W$ est tel que

$$x \in W \subset A^c ;$$

ceci établit que A^c est ouvert, donc A est fermé.

Soit $x_1, x_2, \ldots$ une suite dans A ; posons F_n pour l'adhérence de l'ensemble formé par $x_n, x_{n+1}, \ldots$ Puisque A est fermé, on a que F_n est un ensemble non vide fermé tel que

$$F_n \subset A , \quad F_{n+1} \subset F_n , \quad n = 1, 2, \ldots$$

Démontrons que l'intersection de tous les F_n est non vide. Sinon, les ensembles $V_n = F_n^c$ formeraient un recouvrement ouvert de A (en fait, de $\mathbb{R}^N$) d'où, selon (a), on pourrait extraire un recouvrement fini $V_{n_1}, \ldots, V_{n_k}$ de A, ce qui donnerait, par complémentation, $F_{n_1}, \ldots, F_{n_k}$, tels que

$$F_{n_1} \cap \ldots \cap F_{n_k} \subset A^c ;$$

mais les F_n étant contenus dans A, on aurait alors que l'intersection des F_{n_j} est vide, ce qui est absurde, car les F_n étant une suite d'ensembles non vides et décroissants, l'intersection des F_{n_j} est égale à $F_m \neq \emptyset$ où m est le maximun de $n_1, \ldots, n_k$.

Soit $x \in \bigcap_n F_n$; x est dans A et il adhère à chaque ensemble formé par $x_n, x_{n+1}, \ldots$; ainsi, pour chaque $n = 1, 2, \ldots$, on peut choisir un $x_{k(n)}$ tel que

$$\|x - x_{k(n)}\| < \frac{1}{n} , \quad k(n) \geqslant n ;$$

la sous-suite $\{x_{k(n)}\}$ converge vers $x \in A$.

(b)⇒(c) Il est clair que A est fermé (prop. 1.4.2). Si A ne possède pas la seconde propriété formulée dans (c), on aura un $\varepsilon > 0$ tel que A n'est pas

recouvrable par une famille *finie* de boules de type $B(a;\varepsilon)$, $a \in A$; soit x_1 un point quelconque de A ; par récurrence, on pourra trouver des points x_2, x_3, ... tels que

$$\|x_n - x_j\| > \varepsilon\,, \quad 1 \leqslant j \leqslant (n-1)\,;$$

une telle suite $\{x_n\}$ dans A ne peut contenir aucune sous-suite convergente ce qui contredit (b).

(c)⇒(a) Nous procédons, de nouveau, par un raisonnement par l'absurde. Supposons la validité de (c) mais non celle de (a) ; il existe alors un recouvrement ouvert $\{V_\alpha\}$, $\alpha \in J$, de A d'où l'on ne peut extraire aucun recouvrement fini de A. La validité de (c) entraîne que A est borné ; soit d le diamètre de A. Puisque A peut être recouvert d'une famille finie de boules de types $B(a;d/2)$, $a \in A$, mais qu'aucune sous-famille finie des V_α ne recouvre A, il est clair qu'il y a une boule

$$B_1 = B\,(a_1;d/2)\,, \quad a_1 \in A$$

tel qu'aucune sous-famille finie des V_α ne recouvre

$$A_1 = A \cap B_1\,;$$

par récurence, on obtient une suite de boules

$$B_n = B\,(a_n;d/2^n)\,, \quad a_n \in A\,,$$

et une suite de parties A_n de A, $n = 1, 2, \ldots$ telles que

$$A_n = A_{n-1} \cap B_n\,, \quad \|a_n - a_{n-1}\| \leqslant \frac{d}{2^{n-1}}\,, \quad n \geqslant 2$$

et telles qu'aucune sous-famille finie des V_α ne recouvre une partie A_n. La suite $\{a_n\}$ ainsi définie dans A est une suite de Cauchy car, si $m < n$,

$$\begin{aligned}\|a_m - a_n\| &\leqslant \|a_m - a_{m-1}\| + \cdots + \|a_{n-1} - a_n\| \\ &< \frac{d}{2^m} + \cdots + \frac{d}{2^{n-1}} < \frac{d}{2^{m-1}}\,;\end{aligned}$$

donc $a_n \to a$ $(n \to \infty)$ et $a \in A$ car A est fermé. Or, il existe un $\alpha' \in J$ tel que $a \in V_{\alpha'}$; $V_{\alpha'}$ étant ouvert, il existe un $\varepsilon > 0$ tel que la boule $B(a;\varepsilon)$ est contenue dans $V_{\alpha'}$; mais, si n est assez grand, on aura

$$\|a_n - a\| < \frac{\varepsilon}{2}\,, \quad \frac{d}{2^n} < \frac{\varepsilon}{2}$$

d'où

$$B_n = B\,(a_n;d/2^n) \subset B(a;\varepsilon)$$

car, si $x \in B_n$,

$$\|x - a\| \leqslant \|x - a_n\| + \|a_n - a\| < \frac{d}{2^n} + \frac{\varepsilon}{2} < \varepsilon\,.$$

Ainsi, pour n assez grand,

$$A_n \subset B_n \subset B(a;\varepsilon) \subset V_{\alpha'}$$

ce qui contredit le fait qu'aucune partie A_n n'est recouvrable par une sous-famille finie des V_α.

On a ainsi l'équivalence entre (a), (b) et (c). Aussi l'implication, (c)⇒(d) est évidente.

(d)⇒(a) Soit A un ensemble fermé et borné quelconque dans $\mathbb{R}^N$. Comme A est borné, il existe un $T > 0$ tel que le pavé fermé

$$P = [-T, T] \times \cdots \times [-T, T] \subset \mathbb{R}^N$$

contient A ; montrons d'abord que P satisfait (c). En effet, P est fermé (Ex. 10, § 1.4.7) ; pour un $n \in \mathbb{N}^*$ fixe, posons

$$K = \left\{ -T + \frac{2Tk}{n} : k = 0, \ldots, n \right\} \subset [-T, T]$$

$$S = K \times \cdots \times K \subset P.$$

Il est clair que K et, par conséquent, S sont les ensembles finis ; si $x \in P$, on a évidemment un $y \in S$ tel que

$$|x - y| < \frac{1}{n}$$

car

$$|x - y| = \max_i |x_i - y_i|, \quad x = (x_1, \ldots, x_N),\ y = (y_1, \ldots, y_N)$$

et il suffit de choisir, pour chaque i, un $y_i \in K$ de sorte que

$$|x_i - y_i| < \frac{1}{n}.$$

Comme

$$\|x - y\| < \frac{\sqrt{N}}{n}$$

il suffit de choisir n tel que

$$\frac{\sqrt{N}}{n} < \varepsilon$$

pour avoir

$$P \subset \bigcup \{ B(y;\varepsilon) : y \in S \}.$$

Ainsi P satisfait (c), donc aussi (a) et (b). Si $A \subset P$ est fermé, nous établissons maintenant que A satisfait (b).

Soit $\{x_n\}$, $n \in \mathbb{N}$, une suite dans A, donc aussi dans P ; comme P satisfait (b), il existe une sous-suite de $\{x_n\}$ qui converge vers un point x de P ; comme A est fermé, x doit être forcément dans A. Ainsi, A satisfait (b) et, par voie de conséquence, (a) et (c). ♦

1.7.2 Remarques

La proposition 1.7.1 caractérise complètement les parties compactes de $\mathbb{R}^N$; celles-ci sont exactement les parties fermées et bornées de $\mathbb{R}^N$. Dans certaines considérations, on a besoin d'autres énoncés qui se déduisent facilement de la proposition 1.7.1. Voici quelques exemples.

(1) *Chaque partie infinie, bornée A dans $\mathbb{R}^N$ possède un point d'accumulation.*

En effet, si A est bornée, $\overline{A}$ l'est aussi (conséquence immédiate de la proposition 1.4.2) ; donc $\overline{A}$ est compact (prop. 1.7.1(d)) ; si $\{x_n\}$ est une suite d'éléments distincts dans A, on aura une sous-suite $\{x_{n(k)}\}$, $k = 1, 2, \ldots$, convergeant vers un point x de $\overline{A}$ et x sera alors un point d'accumulation de A.

(2) *Soit $\{x_n\}$ une suite bornée dans $\mathbb{R}^N$; il existe une sous-suite $\{x_{n(k)}\}$, $k = 1, 2, \ldots$, telle que $x_{n(k)} \to x \in \mathbb{R}^N$ $(k \to \infty)$.*

C'est une conséquence immédiate de la proposition 1.7.1(b).

(3) *Soit $\{x_n\}$ une suite bornée dans $\mathbb{R}^N$ telle que toute sous-suite convergente de $\{x_n\}$ converge vers un même élément a de $\mathbb{R}^N$; alors $x_n \to a$ lorsque $n \to \infty$.*

Sinon, on aura un $\varepsilon > 0$ et une suite d'entiers $n(1) < n(2) < \ldots$ telle que

$$\|x_{n(k)} - a\| > \varepsilon\,, \quad k \geqslant 1\,;$$

mais aucune sous-suite convergente de $\{x_{n(k)}\}$ ne peut alors converger vers a ; or, des sous-suites convergentes existent car $\{x_{n(k)}\}$ est une suite bornée.

(4) *Soit $\{F_\alpha\}$, $\alpha \in J$, une famille de parties fermées contenues dans une partie compacte A de $\mathbb{R}^N$ telle que toute intersection d'une sous-famille finie est non vide. Alors*

$$\bigcap_{\alpha \in J} F_\alpha \neq \emptyset\,.$$

En particulier, si $F_1 \supset F_2 \supset \ldots$ est un suite décroissante de parties fermées non vides dans $\mathbb{R}^N$ telle que F_1 est bornée, alors

$$F = \bigcap_n F_n \neq \emptyset\,.$$

En effet, si

$$\bigcap_{\alpha} F_\alpha = \emptyset$$

alors $V_\alpha = F_\alpha^c$, $\alpha \in J$, définissent un recouvrement de $\mathbb{R}^N$, donc de A ; puisque A est compact, il existe une sous-famille finie $V_{\alpha_1}, \ldots, V_{\alpha_n}$ qui recouvre A d'où, par complémentation,

$$F_{\alpha_1} \cap \ldots \cap F_{\alpha_n} \subset A^c ;$$

or, chaque F_α étant contenu dans A, on aura que cette dernière intersection est vide ce qui contredit l'hypothèse de (4).

(5) Un ensemble A de $\mathbb{R}^N$ s'appelle ***précompact*** si, pour tout $\epsilon > 0$, A est contenu dans une réunion finie des boules de type $B(a;\varepsilon)$, $a \in A$ (seconde partie de l'affirmation (c) de la proposition 1.7.1). On peut alors énoncer : *une partie A de $\mathbb{R}^N$ est précompacte si et seulement si A est borné.*

En effet, si A est borné, $\overline{A}$ est fermé et borné, donc compact ; ainsi, il existe des points $x_1, \ldots, x_n$ dans $\overline{A}$ tels que

$$\overline{A} \subset \bigcup_j B(x_j;\varepsilon/2) ;$$

choisissons $a_j \in A$ tel que

$$\|a_j - x_j\| < \frac{\varepsilon}{2}, \quad j = 1, \ldots, n,$$

ce qui est possible car chaque point de $\overline{A}$ adhère à A ; on vérifie immédiatement que

$$\bigcup_j B(a_j;\varepsilon) \supset \overline{A} \supset A .$$

D'autre part, il est évident qu'un ensemble précompact A est borné.

Une autre façon de formuler les énoncés (1) et (2) ci-dessus est d'introduire la définition suivante : *une suite $\{x_n\}$ dans $\mathbb{R}^N$ possède un point d'accumulation x* s'il existe une sous-suite $\{x_{n(k)}\}$, $k = 1, 2, \ldots$ qui converge vers x ; une définition équivalente est que tout voisinage ouvert V de x contient une infinité des termes x_n. En effet, la première définition entraîne la seconde et la réciproque s'établit en prenant V, successivement, comme $B(x;1/k)$, $k = 1, 2, \ldots$ et en sélectionnant $x_{n(k)}$ dans $B(x;1/k)$ tel que $n(k) < n(k+1)$.

On a alors l'énoncé suivant.

(6) *Toute suite bornée dans $\mathbb{R}^N$ possède un point d'accumulation dans $\mathbb{R}^N$.*

La démonstration est laissée comme exercice (cf. la preuve, (a)⇒(b) de la proposition 1.7.1).

A titre d'illustration donnons quelques contre-exemples simples. L'ensemble

$$A =]0, 1] \subset \mathbb{R}$$

n'est pas compact car A n'est pas fermé ; les ensembles

$$V_n = \left]\frac{1}{n}, 2\right[, \quad n = 2, 3, \ldots$$

forment un recouvrement ouvert de A tel qu'aucune sous-famille finie de celui-ci ne soit un recouvrement de A. Aussi, la suite $\{1/n\}$, $n = 1, 2, \ldots$ dans A ne peut pas contenir une sous-suite qui converge vers *un élément* de A car $1/n \to 0$ $(n \to \infty)$ et $0 \notin A$. Prenons

$$A = [0, \infty[\subset \mathbb{R} ;$$

A est fermé mais non borné, donc non compact ; si

$$V_n =]-1, n[, \quad n = 1, 2, \ldots$$

alors les V_n forment un recouvrement ouvert de A ne permettant aucun recouvrement fini de A. Aussi, la suite $\{n\}$, $n = 1, 2, \ldots$, dans A ne peut avoir aucune sous-suite convergente dans A (ou même dans $\mathbb{R}$).

Il est utile peut-être de remarquer que, dans la définition de la compacité, il ne s'agit que des recouvrements *ouverts*. Pour un recouvrement quelconque (ou même recouvrement fermé, c'est-à-dire un recouvrement dont les ensembles sont tous fermés), on ne peut pas avoir des énoncés du type 1.7.1. Par exemple, si

$$A = \left\{0, 1, \frac{1}{2}, \frac{1}{3}, \ldots\right\}$$

alors A est compact mais du recouvrement fermé

$$V_0 = \{0\}, \quad V_n = \left\{\frac{1}{n}\right\}, \quad n = 1, 2, \ldots$$

on ne peut extraire aucun recouvrement fini de A.

Les fonctions continues définies dans un ensemble compact jouissent des propriétés importantes ; quelques unes de celles-ci sont données dans la proposition suivante.

1.7.3 Proposition

Soit $f : A \to \mathbb{R}^M$ une fonction continue définie dans un ensemble compact A contenu dans $\mathbb{R}^N$.

(i) L'ensemble $f(A)$ est compact dans $\mathbb{R}^M$. En particulier, si $M = 1$, f atteint ses bornes dans le sens qu'ils existent ξ, η dans A tels que

$$f(\xi) = \sup_{x \in A} f(x), \quad f(\eta) = \inf_{x \in A} f(x).$$

(ii) f est uniformément continue.

(iii) Si f est injective alors f est un homéomorphisme entre A et $f(A)$.

DÉMONSTRATION

(i) Soit $\{V_\alpha\}$, $\alpha \in J$, un recouvrement ouvert de $f(A)$; pour chaque α, $f^{-1}(V_\alpha)$ est ouvert dans A (car f est continue : proposition 1.6.2(ii)), de sorte que

$$f^{-1}(V_\alpha) = W_\alpha \cap A$$

où W_α est un ensemble ouvert de $\mathbb{R}^N$. On voit immédiatement que $\{W_\alpha\}$, $\alpha \in J$, est un recouvrement ouvert de A ; la compacité de A donne une famille finie $W_{\alpha_1}, \ldots, W_{\alpha_n}$ qui recouvre A ; on en déduit que $V_{\alpha_1}, \ldots, V_{\alpha_n}$ est un recouvrement fini de $f(A)$ extrait du recouvrement ouvert $\{V_\alpha\}$, $\alpha \in J$, de $f(A)$. Ceci établit la compacité de $f(A)$.

Si $M = 1$, $f(A)$ est une partie compacte de $\mathbb{R}$, ce qui veut dire que $f(A)$ est fermé et borné dans $\mathbb{R}$; on a donc

$$\sup_{x\in A} f(x) = L \in f(A), \quad \inf_{x\in A} f(x) = \ell \in f(A)$$

d'où l'existence de ξ, η dans A tels que

$$L = f(\xi), \quad \ell = f(\eta).$$

(ii) Il faut voir que, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que

$$\|f(x) - f(y)\| < \varepsilon \quad \text{si } \|x - y\| < \delta,\ x, y \in A.$$

Sinon, on aura un $\varepsilon > 0$ et des suites $\{x_n\}$, $\{y_n\}$ dans A tels que

$$\|f(x_n) - f(y_n)\| \geqslant \varepsilon, \quad \|x_n - y_n\| < \frac{1}{n}, \quad n = 1, 2, \ldots \tag{1.9}$$

La compacité de A garantit l'existence d'une sous-suite $\{x_{n(k)}\}$ telle que

$$\lim_{k\to\infty} x_{n(k)} = a \in A$$

d'où, on aura, selon la seconde inégalité de (1.9),

$$\lim_{k\to\infty} y_{n(k)} = a$$

et

$$\lim_{k\to\infty} f\big(x_{n(k)}\big) = f(a) = \lim_{k\to\infty} f\big(y_{n(k)}\big)$$

à cause de la continuité de f ; mais cette dernière égalité contredit la première inégalité de (1.9).

(iii) Posons $B = f(A)$; comme A est compact, B l'est aussi. Pour établir que f est un homéomorphisme entre A et B, il suffit de voir que si E est A-fermé alors $f(E)$ est B-fermé (Ex. 6, § 1.6.4). Or, si E est A-fermé, E est compact (utiliser la compacité de A et la section 1.5) d'où $f(E)$ est compact. Mais $f(E)$ est contenu dans B qui est compact ; donc (de la compacité de B et sect. 1.5), $f(E)$ est B-fermé. ♦

1.7.4 Exercices

[1] Si A et B sont deux ensembles compacts dans $\mathbb{R}^N$, vérifier que $A \cup B$, $A \cap B$, $A + B$ sont compacts où $A + B = \{x = a + b : a \in A,\ b \in B\}$.

[2] Soient $A_1, A_2, \ldots$ une suite infinie d'ensembles compacts dans $\mathbb{R}^N$. Vérifier que $\bigcap_n A_n$ est compact mais que $\bigcup_n A_n$ peut être compact ou non.

[3] Soit $A \subset X \subset \mathbb{R}^N$. En suivant la discussion de la section 1.5, définir la notion de X-compacité de A ; démontrer ensuite que A est X-compact si et seulement si A est compact. Vérifier par un exemple que A peut être borné et X-fermé sans être compact.

[4] Soit $f : \mathbb{R}^N \to \mathbb{R}_+$ continue ; si $f(x) \to 0$ lorsque $\|x\| \to \infty$ démontrer que

$$M = \sup_{x \in \mathbb{R}^N} f(x) < \infty\,, \quad \exists\, \xi \in \mathbb{R}^N \text{ tel que } f(\xi) = M\,.$$

[5] Soit p une norme quelconque dans $\mathbb{R}^N$; posons

$$M = \sup_{\|x\|=1} p(x)\,, \quad m = \inf_{\|x\|=1} p(x)\,;$$

vérifier que $0 < m \leqslant M < \infty$; en déduire que

$$m\|x\| \leqslant p(x) \leqslant M\|x\|\,, \quad x \in \mathbb{R}^N\,.$$

Conclure que si q est une autre norme dans $\mathbb{R}^N$ alors ils existent c, C, $0 < c \leqslant C < \infty$, tels que

$$cq(x) \leqslant p(x) \leqslant Cq(x)\,, \quad x \in \mathbb{R}^N\,.$$

[6] Soit $f : J_1 \times J_2 \to \mathbb{R}$ continue où J_1, J_2 sont deux intervalles compacts de $\mathbb{R}$; posons

$$\varphi(x) = \sup_{y \in J_2} f(x, y)\,.$$

Démontrer que $\varphi : J_1 \to \mathbb{R}$ est continue.

7 Soit A un ensemble non compact dans $\mathbb{R}$. Vérifier que :

(i) il existe $f \in C(A)$ qui n'est pas bornée ;
(ii) il existe $f \in C(A)$ qui est bornée mais que $\sup_x f(x)$ n'est pas atteint dans A ;
(iii) si A est borné, il existe $f \in C(A)$ qui n'est pas uniformément continue.

(Essayer les fonctions de type $(x-a)^{-1}$, $\{1+(x-a)^2\}^{-1}$, etc.) L'ensemble $\mathbb{Z}$ est non compact mais toute fonction de $C(\mathbb{Z})$ est uniformément continue.

8 Démontrer qu'une fonction uniformément continue $f : \mathbb{R} \to \mathbb{R}$ satisfait une inégalité

$$|f(x)| \leqslant \alpha|x| + \beta\,, \quad x \in \mathbb{R}$$

pour α, $\beta \geqslant 0$ appropriés.

9 Soit A une partie bornée de $\mathbb{R}^N$. Vérifier que $\overline{A}$ est aussi borné donc compact.

1.8 CONNEXITÉ

La notion de connexité d'un ensemble dans $\mathbb{R}^N$ provient de l'idée intuitive que celui-ci est « d'un seul morceau » ou « d'un seul tenant ». Ainsi, un intervalle, une boule, une sphère sont perçus comme étant « d'un seul morceau », alors que la réunion de deux boules ouvertes sans points communs apparaissent évidemment « non connexe ». Il y a plusieurs façons mathématiques de formaliser la notion intuitive de connexité dont deux sont formulées dans ce qui suit.

1.8.1 Définitions

Soit $A \subset \mathbb{R}^N$; A s'appelle ***connexe*** si les seuls sous-ensembles de A qui sont à la fois ouverts et fermés dans A (sect. 1.5) sont les ensembles $\emptyset$ et A. Une définition équivalente est que A est connexe si A *ne peut pas être représenté comme*

$$A = A_1 \cup A_2\,, \quad A_1 \cap A_2 = \emptyset$$

où A_1, A_2 sont deux ensembles non vides ouverts (fermés) dans A. Autrement dit, si A est représenté comme la réunion de deux ensembles disjoints, ouverts dans A (ou fermés dans A), alors un des deux est vide.

Un ***chemin dans*** $A \subset \mathbb{R}^N$ est une application continue

$$\varphi : [a\,, b] \to A$$

où $[a\,, b]$ est un intervalle compact dans $\mathbb{R}$; on dit que ce chemin ***relie*** les points $\varphi(a)$ et $\varphi(b)$ dans A. L'ensemble A s'appelle ***connexe par arcs*** si deux

points quelconques x, y de A sont reliés par un chemin dans A. Un chemin $\varphi : [a, b] \to \mathbb{R}^N$ s'appelle ***simple*** si φ est une application injective ; φ s'appelle ***polygonal*** s'ils existent

$$t_0 = a < t_1 < \ldots < t_n = b$$

tels que φ est une application affine dans l'intervalle $[t_{j-1}, t_j]$, $j = 1, \ldots, n$ (fig. 1.3) ; on dira que φ est obtenu en ***juxtaposant*** les segments $[x_{j-1}, x_j]$, $x_j = \varphi(t_j)$.

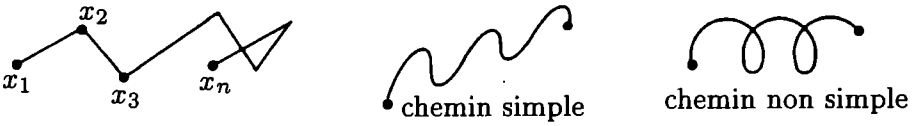


Fig. 1.3

Soient $A \subset \mathbb{R}^N$, X un ensemble quelconque, $f : A \to X$. On dit que f est ***localement constante*** si tout point a de A possède un voisinage ouvert V tel que f est constante dans $V \cap A$.

Il est clair que l'ensemble vide est aussi bien connexe que connexe par arcs. La proposition suivante résume les propriétés importantes des ensembles connexes et ceux connexes par arcs.

1.8.2 Proposition

(i) Les seuls ensembles non vides connexes dans $\mathbb{R}$ sont les intervalles ou les ensembles réduits à un point.

(ii) Si $f \in C(A; \mathbb{R}^M)$ et A est connexe (connexe par arcs) dans $\mathbb{R}^N$, alors $f(A)$ est connexe (connexe par arcs) dans $\mathbb{R}^M$.

(iii) (Théorème des valeurs intermédiaires) Si $f \in C(A)$ et A est un sous-ensemble connexe non vide de $\mathbb{R}^N$, alors $f(A)$ est un intervalle ou un sous-ensemble réduit à un point de $\mathbb{R}$; en particulier, si f prend, dans A, les valeurs α et β, alors f assume, dans A, toute valeur intermédiaire entre α et β.

(iv) Soit $A \subset \mathbb{R}^N$. Si A est connexe par arcs alors A est connexe. Si A est ouvert et connexe alors A est connexe par arcs ; dans ce cas, deux points de A peuvent être reliés par un chemin polygonal, simple, dans A.

(v) Soit $A \subset \mathbb{R}^N$, A connexe. Si $f : A \to X$ est une application localement constante, X étant un ensemble quelconque, alors f est constante.

DÉMONSTRATION

(i) Soit A une partie non vide connexe de $\mathbb{R}$ contenant plus d'un point. Si $x < y$ sont deux points de A, montrons d'abord que l'intervalle fermé $[x, y]$ est contenu dans A ; en effet, sinon, on aurait un z tel que

$$z \notin A, \quad x < z < y ;$$

mais alors,

$$A_1 = A \cap]-\infty\,,\, z[= A \cap]-\infty\,,\, z]$$

serait un ensemble non vide, à la fois ouvert et fermé dans A, sans qu'il soit égal à A ce qui contredirait la connexité de A. On conclut que si

$$\alpha = \inf_{x\in A} x\,, \quad \beta = \sup_{x\in A} x$$

alors A est égal à un intervalle dont les extrémités sont α et β. (Par exemple, si $\alpha \in A$, $\beta \notin A$, on aura $A = [\alpha\,,\, \beta[$, etc.)

Si $A = \{a\}$, A est connexe. Démontrons qu'un intervalle quelconque A dans $\mathbb{R}$ d'extrémités $\alpha < \beta$ (finies ou non) est connexe. Sinon, on pourrait écrire A comme la réunion de deux ensembles, non vides, disjoints, A_1, A_2, chaque A_i étant ouvert dans A. La propriété des A_i qui va nous servir est la suivante : soit $t \in A_i$; si $t \neq \beta$, il existe $t' > t$ tel que $]t\,,\, t'[$ est contenu dans A_i ; si $t \neq \alpha$, il existe $t' < t$ tel que $]t'\,,\, t[$ est contenu dans A_i. C'est une conséquence du fait que les A_i sont ouverts dans A qui est un intervalle. Sans perte de généralité, nous pouvons supposer qu'il existe $x_1 \in A_1$, $x_2 \in A_2$, $x_1 < x_2$; posons

$$t = \sup\{x : x \in A_1 \cap [x_1\,,\, x_2]\}\ ;$$

il est clair que t est dans l'intervalle A ; or t ne peut appartenir ni à A_1 ni à A_2, car si $t \in A_1$ on aurait un intervalle $]t\,,\, t'[$ contenu dans A_1 contredisant la définition de t, et si $t \in A_2$ on aurait un intervalle $]t'\,,\, t[$ contenu dans A_2 contredisant, de nouveau, la définition de t ; au total, t est dans A mais ni dans A_1, ni dans A_2, une absurdité. On conclut que l'intervalle A doit être connexe.

(ii) Soit $f \in C(A; \mathbb{R}^M)$; supposons d'abord que A est une partie connexe de $\mathbb{R}^N$; posons $B = f(A)$. Si E est à la fois ouvert et fermé dans B, $f^{-1}(E)$ sera aussi à la fois ouvert et fermé dans A (prop. 1.6.2(ii)), d'où, selon la connexité de A, $f^{-1}(E)$ sera égal à $\emptyset$ ou A, ce qui donne que E est ou bien vide ou bien B ; ainsi, B est connexe. Supposons maitenant que A est connexe par arcs ; si u, v sont dans $B = f(A)$ et x, y sont dans A tels que

$$u = f(x)\,, \quad v = f(y)$$

prenons un chemin φ dans A reliant x à y, c'est-à-dire

$$\varphi : [a\,,\, b] \to A\,, \quad \varphi(a) = x\,,\ \varphi(b) = y\,,\ \varphi \text{ continue}\ ;$$

alors $f \circ \varphi$ sera un chemin dans B qui relie u à v ce qui démontre la connexité par arcs de B.

(iii) Si $f \in C(A)$ et A est connexe, alors $f(A)$ sera connexe dans $\mathbb{R}$ (selon (ii)) d'où le résultat selon (i).

(iv) Soit A une partie connexe par arcs de $\mathbb{R}^N$; pour voir que A est connexe, nous montrons l'impossibilité d'écrire A comme la réunion de A_1, A_2, où les A_i sont non vides, disjoints et fermés dans A. Soit $x_i \in A_i$, $i = 1, 2$ et soit φ un chemin dans A qui relie x_1 à x_2 dans A :

$$\varphi : [a, b] \to A, \quad \varphi(a) = x_1, \ \varphi(b) = x_2 .$$

Soit T la borne supérieure des t dans $[a, b]$ tels que $\varphi(t)$ est dans A_1 ; il est évident que

$$a \leqslant T < b .$$

Si $y = \varphi(T)$, nous allons voir que y doit appartenir aux deux ensembles A_1 et A_2, ce qui contredit le fait qu'ils sont disjoints. En effet, si $T = a$, $y = x_1$ et $y \in A_1$; si $T > a$, y est la limite d'une suite d'éléments $\varphi(t_n)$ de A_1 où

$$a < t_n < T, \quad t_n \to T \ (n \to \infty),$$

d'où $y \in A_1$ car A_1 est fermé dans A. De même, $y \in A_2$ car on a une suite $\{t_n\}$ telle que

$$T < t_n < b, \quad t_n \to T \ (n \to \infty), \quad \varphi(t_n) \in A_2$$

d'où, par la continuité de φ, $\varphi(t_n) \to \varphi(T) = y$ $(n \to \infty)$ et $y \in A_2$ car A_2 est fermé dans A. On conclut que A doit être connexe.

Soit A une partie non vide, ouverte et connexe dans $\mathbb{R}^N$; nous devons établir que deux points quelconques de A peuvent être reliés dans A par un chemin polygonal simple. Fixons un point x de A et appelons S le sous-ensemble formé des points y de A qui peuvent être reliés à x par un chemin polygonal simple, dans A. Il est clair que S est non vide car toute boule $B(x; r)$ contenue dans A est constituée de points de S ; nous allons voir que S est à la fois ouvert et fermé dans A, d'où l'on conclut que $A = S$.

D'abord, S est ouvert (ce qui est équivalent à « S est ouvert dans A » car A est ouvert) : en effet, soit $y \in S$; soit φ un chemin polygonal simple reliant x à y dans A, obtenu en juxtaposant les segments $[x_{j-1}, x_j]$, $j = 1, \ldots, n$, $x_0 = x$, $x_n = y$. Choisissons un $\varepsilon > 0$ tel que $B(y; \varepsilon)$ est contenu dans A et tel que $B(y; \varepsilon)$ intersecte seulement le segment $[x_{n-1}, x_n]$ et n'intersecte pas les autres segments $[x_{j-1}, x_j]$, $j < n$ (s'il y en a) ; il est clair que tout point z de $B(y; \varepsilon)$ peut être relié à x par un chemin polygonal simple, en prolongeant un chemin φ donné par un segment $[y, z]$ (si nécessaire). Ainsi, $B(y; \varepsilon) \subset S$ et on conclut que S est ouvert. Pour établir que S est fermé dans A, nous partons d'un y, un point de A qui adhère à S, pour voir que $y \in S$. Comme A est ouvert, on peut choisir un $r > 0$ tel que $B(y; r)$ est contenu dans A ; puisque y adhère à S, il existe un z dans

$$S \cap B(y; r) ;$$

on a alors un chemin polygonal simple φ, reliant x à z dans A, obtenu en juxtaposant les segments $[x_{j-1}, x_j]$, $j = 1, \ldots, n$, $x_0 = x$, $x_n = z$; si y

ne se trouve pas déjà sur ce chemin φ (cas où y est évidemment dans S), on peut prolonger le chemin φ en un autre (noté φ_1) en ajoutant le segment $[z, y]$ à φ. Le chemin φ_1 ainsi obtenu est clairement polygonal ; si φ_1 est simple, la démonstration que $y \in S$ est achevée ; sinon le segment $[z, y]$ doit intersecter un des segments $[x_{j-1}, x_j]$, $j = 1, \ldots, n$; soit $k \leqslant n$ tel que

$$[z, y] \cap [x_{j-1}, x_j] = \emptyset, \quad j = 1, \ldots, k-1$$

mais que

$$[z, y] \cap [x_{k-1}, x_k]$$

est non vide, donc contient exactement un point p ; alors le chemin obtenu par la juxtaposition des segments

$$[x_0, x_1], \ldots, [x_{k-1}, p], [p, y]$$

donne un chemin polygonal simple reliant $x_0 = x$ à y dans A. Donc $y \in S$.

(v) Si $a \in A$, posons

$$E_a = \{x \in A : f(x) = f(a)\} ;$$

il est clair que E_a est non vide (car $a \in E_a$) ; de plus, E_a est ouvert dans A car, si $x \in E_a$, il existe un voisinage ouvert V de x tel que f est constamment égal à $f(x) = f(a)$ dans $V \cap A$ de sorte que

$$V \cap A \subset E_a$$

d'où la conclusion que E_a est ouvert dans A. Notons que si $f(a) \neq f(a')$ alors E_a, $E_{a'}$ sont disjoints. Fixons un a ; si $E_a \neq A$ on aura une partition de A en A_1, A_2 où

$$A_1 = E_a, \quad A_2 = \bigcup \{E_{a'} : a' \text{ tel que } f(a') \neq f(a)\}$$

avec A_1, A_2 disjoints et ouverts dans A (noter que A_2 est ouvert dans A, car il est une réunion des $E_{a'}$ ouverts dans A), ce qui contredit la connexité de A. On conclut que f est constante dans A. ♦

1.8.3 Remarques sur la connexité

(1) *Si chaque A_i, $i \in J$, est connexe dans $\mathbb{R}^N$ et $\bigcap_i A_i$ est non vide, alors*

$$A = \bigcup_{i \in J} A_i$$

est connexe.

En effet, soit $a \in A_i$, $i \in J$; si A n'est pas connexe, on pourra écrire

$$A = B \cup C$$

ou B, C sont non vides, disjoints et ouverts dans A ; supposons, par exemple, que $a \in B$; puisque C est non vide, on a un $\alpha \in J$ tel que $C \cap A_\alpha$ est non vide ; alors

$$B \cap A_\alpha\,, \quad C \cap A_\alpha$$

sont deux ensembles non vides, disjoints, ouverts dans A_α dont la réunion est A_α ce qui contredit la connexité de A_α. ♦

(2) *Si A est une partie connexe de $\mathbb{R}^N$ et*

$$A \subset B \subset \overline{A}$$

alors B est connexe ; en particulier, $\overline{A}$ est connexe.

En effet, si B_1, B_2 est une partition de B en deux ensemble non vides, ouverts dans B, alors $A \cap B_1$, $A \cap B_2$ seront disjoints, ouverts dans A et leur réunion est A ; aussi, les deux ensembles $A \cap B_i$, $i = 1, 2$, sont non vides ; par exemple, si

$$B_1 = V \cap B\,, \quad V \text{ ouvert}\,, \; x \in B_1 \subset B \subset \overline{A}$$

on aura un y tel que

$$y \in V \cap A \subset V \cap B = B_1$$

car $x \in V$ et x adhère à A. Ceci contredit la connexité de A. ♦

(3) *Soit A une partie connexe de $\mathbb{R}^N$; si B est un ensemble dans $\mathbb{R}^N$ tel que*

$$A \cap B\,, \quad A \cap B^c$$

sont non vides alors $A \cap \operatorname{Fr} B$ est non vide.

En effet, si $A \cap \operatorname{Fr} B$ est vide, on aurait

$$A_1 = A \cap \mathring{B} = A \cap B\,, \quad A_2 = A \cap (\operatorname{ext} B) = A \cap B^c$$

et A serait la réunion de A_1, A_2 non vides, disjoints, ouverts dans A ce qui contredit la connexité de A. ♦

Une application intéressante de (3) est la suivante : *si un chemin relie un point $x \in B$ à un point $y \in B^c$ alors le chemin rencontre* $\operatorname{Fr} B$ (où B est une partie quelconque de $\mathbb{R}^N$). En effet, si A est la « trajectoire » du chemin reliant x à y dans $\mathbb{R}^N$, c'est-à-dire

$$A = \varphi([a\,, b])\,, \quad \varphi : [a\,, b] \to \mathbb{R}^N\,, \; \varphi(a) = x\,, \; \varphi(b) = y\,, \; \varphi \text{ continu}\,,$$

alors A étant connexe (car A est l'image par φ, continue, de l'ensemble connexe $[a\,, b]$), l'affirmation résulte immédiatement de (3). Une formulation pittoresque de cette dernière affirmation est la suivante : *un chemin qui ne reste pas*

toujours dans le même pays B doit traverser sa frontière. Cette formulation a donné le surnom *théorème du passage des douanes* à l'énoncé (3).

(4) Soit A un sous-ensemble de $\mathbb{R}^N$; si $a \in A$, on peut considérer la réunion, $C(a)$, de toutes les parties connexes de A qui contiennent le point a ; en se rappelant que $\{a\}$ est une partie connexe de A, on voit, grâce à l'énoncé (1) que $C(a)$ est un ensemble connexe contenu dans A et contenant a. On dit que $C(a)$ est la ***composante connexe*** de a dans A ; c'est le plus grand ensemble connexe contenant a qui est contenu dans A ; il est clair que si A est connexe, $C(a) = A$ pour tout $a \in A$ et d'autre part, si pour un $a \in A$, $C(a) = A$, alors A est connexe. Notons aussi que

$$\text{si } b \in C(a)\,, \quad \text{alors } C(a) = C(b)\,,$$

$$\text{si } b \notin C(a)\,, \quad \text{alors } C(a) \cap C(b) = \emptyset\,.$$

Ainsi, l'ensemble A est la réunion d'une famille de sous-ensembles A_i, $i \in J$, deux à deux disjoints tels que chaque A_i est la composante connexe d'un point de A ; les A_i s'appellent les ***composantes connexes*** de A.

(5) Prenons dans la discussion précédente un ensemble A *ouvert* dans $\mathbb{R}^N$; alors, les composantes connexes A_i sont ouvertes aussi ; en effet, si $a \in A_i$, il existe une boule $B(a;\varepsilon)$ contenue dans A ; or, une boule ouverte est connexe par arcs, car deux points quelconques de la boule se relient par un segment dans la boule : ainsi, $B(a;\varepsilon)$ est contenue dans A_i et l'ensemble A_i est ouvert. On conclut par l'énoncé suivant : *chaque ensemble ouvert non vide de $\mathbb{R}^N$ est la réunion d'une famille d'ensembles ouverts et connexes deux à deux disjoints.*

(6) Un ensemble non vide, ouvert et connexe dans $\mathbb{R}^N$ s'appellera un ***domaine*** dans $\mathbb{R}^N$. L'énoncé de (5) peut être alors formulé comme suit : *chaque ouvert non vide de $\mathbb{R}^N$ est la réunion des domaines deux à deux disjoints.* Ceci permet souvent de restreindre les problèmes étudiés dans un ensemble ouvert à ceux dans un domaine.

1.8.4 Exercices

[1] Un ensemble convexe est connexe ; en particulier, les pavés (ouverts ou fermés ou pas) et les boules (ouvertes ou fermées) sont connexes.

[2] L'image dans $\mathbb{R}^N$ d'un intervalle de $\mathbb{R}$ par une fonction continue est souvent appelée une *courbe* ; une courbe est un ensemble connexe.

[3] Un ensemble connexe non vide dans $\mathbb{R}^N$, non réduit à un point, contient une infinité non dénombrable de points. (Si A est cet ensemble et a, b deux points de A, la fonction f donnée par $f(x) = \|x - a\|$, $x \in A$, est continue et $f(A)$ contient l'intervalle $[\,0\,,\|b - a\|\,]$.)

[4] Démontrer directement qu'un ensemble fini non vide non réduit à un point est non connexe.

5 Un ensemble dénombrable quelconque dans $\mathbb{R}^N$ est non connexe ; ses composantes connexes sont les ensembles de type $\{x\}$.

6 Une sphère est un ensemble connexe.

7 L'ensemble $\mathbb{R}^2 \setminus \{(0,0)\}$ est connexe dans $\mathbb{R}^2$ mais $\mathbb{R}^1 \setminus \{0\}$ est non connexe.

8 Soit A un ensemble dans $\mathbb{R}^N$; toute fonction continue $f : A \to \mathbb{Z}$ est constante si et seulement si A est connexe.

9 Soit p un polynôme réel d'une variable réelle de degré impair. Démontrer que $p(x) = 0$ pour au moins un $x \in \mathbb{R}$. (Vérifier que les signes de $p(x)$ changent lorsque l'on passe de x grands et positifs aux x grands et négatifs.)

10 Soit $S = \{x \in \mathbb{R}^N : \|x\| = 1\}$; si $f : S \to \mathbb{R}$ est une fonction continue, démontrer qu'il existe un point $p \in S$ tel que $f(p) = f(-p)$. (Considérer $g(x) = f(x) - f(-x)$; si $g(x) \neq 0$ pour tout $x \in S$, S serait la réunion des ensembles S-ouverts $\{x : g(x) > 0\}$, $\{x : g(x) < 0\}$ non vides, disjoints, ce qui contredit l'exercice 6.)

11 Dans $\mathbb{R}^2$, posons

$$A = \left\{(x,y) : y = \sin\frac{1}{x},\ x > 0\right\}, \quad B = A \cup \{(0,y) : -1 \leqslant y \leqslant 1\}.$$

Voir que A est connexe par arcs (Ex. 2) et que $B = \overline{A}$: en déduire que B est connexe ; mais B est non connexe par arcs, car un point de A ne peut pas être relié à un point de type $(0,y)$ de B par un chemin dans B.

1.9 REMARQUES SUPPLÉMENTAIRES SUR LES FONCTIONS CONTINUES

1.9.1 La fonction $x \mapsto d(x,A)$

Soit A une partie non vide de $\mathbb{R}^N$; posons

$$d(x,A) = \inf_{y \in A} \|x - y\|, \quad x \in \mathbb{R}^N.$$

$d(x,A)$ s'appelle la ***distance du point x à l'ensemble*** A. On a l'inégalité suivante pour $d(x,A)$:

$$|d(x,A) - d(y,A)| \leqslant \|x - y\|, \quad x, y \in \mathbb{R}^N. \tag{1.10}$$

En effet,

$$d(x,A) = \inf_{z\in A} \|x-z\| \leqslant \inf_{z\in A} \{\|x-y\| + \|y-z\|\}$$
$$\leqslant \|x-y\| + \inf_{z\in A} \|y-z\| = \|x-y\| + d(y,A)$$

et, de même,

$$d(y,A) \leqslant \|x-y\| + d(x,A)$$

d'où l'inégalité (1.10).

Une conséquence importante de cette inégalité est que la fonction $x \mapsto d(x,A)$ *est une fonction uniformément continue* ; de plus, on voit que

$$\{x : d(x,A) = 0\} = \overline{A}\,.$$

Si A est une partie non vide de $\mathbb{R}^N$, un ***voisinage ouvert*** de A est un ensemble ouvert contenant A.

Pour chaque $t > 0$, l'ensemble

$$V_t(A) = \{x : d(x,A) < t\}$$

est un voisinage ouvert de A ou $\overline{A}$.

En effet, $x \mapsto d(x,A)$ étant continue, l'ensemble $V_t(A)$ est ouvert ; de ce qui précède, $V_t(A)$ contient $\overline{A}$.

Notons que *si A est un ensemble, non vide, borné dans $\mathbb{R}^N$ alors $V_t(A)$ est borné aussi.* En effet, si δ est le diamètre de A, on vérifie que, pour x, y dans $V_t(A)$,

$$\text{si } x \in A\,,\ y \in A \quad \text{alors } \|x-y\| \leqslant \delta$$
$$\text{si } x \notin A\,,\ y \in A \quad \text{alors } \|x-y\| \leqslant d(x,A) + \delta < t + \delta$$
$$\text{si } x \notin A\,,\ y \notin A \quad \text{alors } \|x-y\| \leqslant \|x-a\| + \|a-y\| \ (\text{si } a \in A)$$
$$< 2(t+\delta)$$

d'où l'on voit que le diamètre de $V_t(A)$ est fini.

Soit A, B deux sous-ensembles non vides de $\mathbb{R}^N$; on pose

$$d(A,B) = \inf_{x\in A} d(x,B) = \inf_{y\in B} d(A,y) = \inf_{\substack{x\in A \\ y\in B}} \|x-y\|\,.$$

Il est clair que

$$d(A,B) = d(B,A)\,, \quad d(A,B) = 0 \text{ si } A \cap B \neq \emptyset\,.$$

Si A est compact, B est fermé et $A \cap B = \emptyset$, alors

$$d(A,B) > 0\,.$$

En effet, la fonction $x \mapsto d(x, B)$ étant continue, elle atteint son infinimum dans l'ensemble compact A ; donc, il existe $a \in A$, tel que

$$d(A, B) = d(a, B) ;$$

si $d(A, B) = 0$ on aurait $d(a, B) = 0$ ce qui entraînerait que $a \in \overline{B} = B$ et a serait à la fois dans A et B contredisant l'hypothèse que A, B sont disjoints.

Si V est un voisinage ouvert d'un ensemble compact non vide A dans $\mathbb{R}^N$, alors il existe un $t > 0$ tel que

$$V_t(A) \subset \overline{V_t}(A) \subset V$$

où

$$V_t(A) = \{x : d(x, A) < t\} .$$

En effet, soit $\delta = d(A, V^c)$; A étant compact, V^c fermé et disjoint de A, on a que $\delta > 0$; si $0 < t < \delta$, on aura

$$V_t(A) \subset \overline{V_t}(A) \subset \{x : d(x, A) \leqslant t\} \subset V$$

à cause du fait que $\{x : d(x, A) \leqslant t\}$ est un ensemble fermé (car $x \mapsto d(x, A)$ est continue) et que $x \in V^c$ entraîne que

$$d(x, A) \geqslant \delta$$

de sorte que si $d(x, A) \leqslant t$, alors $d(x, A) < \delta$ d'où $x \in V$. Les fonctions $x \mapsto d(x, A)$ jouent un rôle théorique très utile. Par exemple, on peut établir *le lemme* dit *de Urysohn* dans $\mathbb{R}^N$ en utilisant ces fonctions : *si A, B sont deux ensembles fermés, non vides, disjoints dans $\mathbb{R}^N$, alors il existe une fonction $f \in C(\mathbb{R}^N)$ telle que*

$$0 \leqslant f \leqslant 1 , \quad A = f^{-1}(0) , \quad B = f^{-1}(1) .$$

Il suffit de définir

$$f(x) = \frac{d(x, A)}{d(x, A) + d(x, B)} .$$

Notons aussi que si

$$V = \left\{x : f(x) < \frac{1}{2}\right\} , \quad W = \left\{x : f(x) > \frac{1}{2}\right\}$$

alors V, W sont deux ensembles ouverts disjoints tels que

$$A \subset V , \quad B \subset W .$$

Donc *deux ensembles fermés disjoints dans $\mathbb{R}^N$ peuvent être séparés par deux ensembles ouverts disjoints.*

En posant

$$g(x) = a + f(x)(b - a) \quad (a < b)$$

dans ce qui précède, on obtient $g \in C(\mathbb{R}^N)$ telle que

$$g(x) = a \text{ si } x \in A\,, \quad g(x) = b \text{ si } x \in B\,, \quad a \leqslant g(x) \leqslant b\,,\ x \in \mathbb{R}^N\,.$$

1.9.2 Limites des fonctions continues

Les exemples simples montrent qu'une suite de fonctions continues peuvent converger vers une fonction discontinue. Ainsi, si

$$f_n(x) = x^n\,, \quad 0 \leqslant x \leqslant 1$$

alors

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{si } 0 \leqslant x < 1 \\ 1 & \text{si } x = 1. \end{cases}$$

Ici, $f_n \in C([0\,,\,1])$, $f_n(x) \to f(x)$ $(n \to \infty)$ pour tout $x \in [0\,,\,1]$, mais f n'est pas continue au point $x = 1$.

Il existe une condition suffisante classique sur la nature de la convergence de f_n vers f qui garantit que la fonction limite f sera continue aussi.

Soient $f_n : X \to \mathbb{R}$, $n \in \mathbb{N}$, X un ensemble quelconque. On dit que f_n ***converge uniformément*** (dans X) vers une fonction f dans X si

$$\sup_{x\in X} |f_n(x) - f(x)| \to 0 \quad \text{lorsque } n \to \infty\,.$$

On peut exprimer cette condition de diverses manières :

(i) pour tout $\varepsilon > 0$, il existe un n_0 (dépendant de ε mais indépendant de $x \in X$) tel que

$$|f_n(x) - f(x)| < \varepsilon\,, \quad x \in X\,,\ n \geqslant n_0\ ;$$

(ii) pour tout $\varepsilon > 0$, il existe un n_0 (dépendant de ε mais indépendant de $x \in X$) tel que

$$|f_m(x) - f_n(x)| < \varepsilon\,, \quad x \in X\,,\ m, n \geqslant n_0\ ;$$

(iii) $\sup_{x\in X} |f_m(x) - f_n(x)| \to 0$, lorsque $m, n \to \infty$.

L'avantage des formulations (ii) et (iii) est qu'elles évitent la mention de la fonction limite f dont l'existence est une conséquence de la propriété de complétion des nombres réels.

On a alors la proposition suivante : *si $f_n \in C(X)$, $X \subset \mathbb{R}^N$, $n = 1, 2, \ldots$, est une suite de fonctions qui converge uniformément alors la fonction limite $f \in C(X)$.*

En effet, si $a \in X$,

$$|f(x) - f(a)| \leqslant |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| ;$$

pour $\varepsilon > 0$ donné, choisir d'abord n_0 (grâce à la convergence uniforme de f_n vers f) tel que

$$\sup_{y \in X} |f_n(y) - f(y)| < \frac{\varepsilon}{3}, \quad n \geqslant n_0$$

et $\delta > 0$ ensuite (grâce à la continuité de f_{n_0} au point a) tel que

$$|f_{n_0}(x) - f_{n_0}(a)| < \frac{\varepsilon}{3}, \quad \|x - a\| < \delta, \ x \in X ;$$

alors

$$|f(x) - f(a)| < \varepsilon \quad \text{si } x \in X, \ \|x - a\| < \delta .$$

Notons comme corollaire la continuité des fonctions représentées par des séries de fonctions continues *uniformément convergentes* : *si*

$$f(x) = \sum_{n=0}^{\infty} f_n(x), \quad x \in X \subset \mathbb{R}^N ,$$

où $f_n \in C(X)$ *et la série converge uniformément dans* X (*dans le sens que* $\{\sum_{n=0}^{T} f_n\}_{T \geqslant 0}$ *converge uniformément vers* f *dans* X), *alors* $f \in C(X)$. C'est une conséquence immédiate de la proposition précédente car la somme finie

$$\sum_{n=0}^{T} f_n$$

est une fonction continue pour $T = 0, 1, \ldots$

Dans les applications, il est commode d'introduire une condition plus restrictive mais plus aisée à vérifier sur une série de fonctions $\sum_n f_n$.

On dit que la série de fonctions réelles

$$\sum_{n=0}^{\infty} f_n$$

dans un ensemble $X \subset \mathbb{R}^N$ est ***normalement convergente*** dans X si

$$\sup_{x \in X} |f_n(x)| \leqslant \alpha_n , \quad \sum_{n=0}^{\infty} \alpha_n < \infty .$$

On voit immédiatement ce qui suit : *Une série* $\sum_n f_n$ *de fonctions réelles normalement convergente dans* $X \subset \mathbb{R}^N$ *est absolument convergente à chaque point* $x \in X$; *elle est aussi uniformément convergente dans* X *vers une fonction réelle* f *dans* X ; *si* $f_n \in C(X)$, $n \in \mathbb{N}$, *alors* $f \in C(X)$.

1.9.3 Prolongement des fonctions continues

Soit $f : X \to Y$; une question théorique importante est la possibilité de prolonger la définition de f à un ensemble plus grand, tout en préservant certaines propriétés de départ de f. Ici, nous considérons le cas où $Y = \mathbb{R}$, $X \subset \mathbb{R}^N$ et f est continue.

Si $X =]0, 1]$, les exemples

$$f(x) = \frac{1}{x} \quad \text{ou} \quad f(x) = \sin \frac{1}{x}$$

montrent qu'il n'est pas toujours possible de prolonger f à l'espace entier, tout en préservant la continuité de f. Mais c'est possible si X est une partie *fermée* dans $\mathbb{R}^N$. Plus précisément, on a la proposition suivante (appelée *théorème de prolongement de Tietze-Urysohn*) : *soit $f \in C(X)$, X fermé dans $\mathbb{R}^N$; il existe $\widetilde{f} \in C(\mathbb{R}^N)$ tel que $\widetilde{f}(x) = f(x)$, $x \in X$; si f est bornée alors $\widetilde{f}$ peut être défini de telle manière que*

$$\sup_{x \in X} |f(x)| = \sup_{x \in \mathbb{R}^N} |\widetilde{f}(x)| .$$

DÉMONSTRATION

Nous considérons d'abord le cas de f bornée.

On peut évidemment écarter le cas trivial $f \equiv 0$; posons

$$f_0 = f , \quad M_0 = \sup_{x \in X} |f(x)| ,$$

$$A_0 = \left\{ x \in X : f_0(x) \leqslant -\frac{M_0}{3} \right\} , \quad B_0 = \left\{ x \in X : f_0(x) \geqslant \frac{M_0}{3} \right\} .$$

Les ensembles A_0, B_0 sont disjoints et fermés dans X donc fermés (car X est fermé) ; du lemme de Urysohn (§ 1.9.1) on déduit l'existence d'une fonction $g_0 \in C(\mathbb{R}^N)$ telle que

$$g_0(x) = -\frac{M_0}{3} , \ x \in A_0 , \quad g_0(x) = \frac{M_0}{3} , \ x \in B_0 , \quad |g_0(x)| \leqslant \frac{M_0}{3} , \ x \in \mathbb{R}^N ;$$

on vérifie que

$$|f_0(x) - g_0(x)| \leqslant \frac{2}{3} M_0 , \quad x \in X .$$

Posons

$$f_1 = f_0 - g_0 , \quad M_1 = \sup_{x \in X} |f_1(x)| \leqslant \frac{2}{3} M_0 .$$

En appliquant le procédé précédent sur f_1, M_1 en place de f_0, M_0 et en continuant récursivement, on obtient une suite de fonctions $g_0, g_1, \ldots$ dans $C(\mathbb{R}^N)$ telles que

$$\left| f(x) - \sum_{i=0}^{n} g_i(x) \right| \leqslant \left(\frac{2}{3}\right)^{n+1} M_0 , \quad x \in X , \; n \geqslant 0 ,$$

$$\sup_{x \in \mathbb{R}^N} |g_n(x)| \leqslant \frac{1}{3} \left(\frac{2}{3}\right)^n M_0 .$$

Posons

$$\widetilde{f}(x) = \sum_{n=0}^{\infty} g_n(x) , \quad x \in \mathbb{R}^N .$$

Selon la définition des fonctions $\{g_n\}$, la série $\sum_n g_n$ est normalement convergente d'où, selons le paragraphe 1.9.2, $\widetilde{f} \in C(\mathbb{R}^N)$; de plus, si $x \in X$,

$$|f(x) - \widetilde{f}(x)| = \lim_{n \to \infty} \left| f(x) - \sum_{i=1}^{n} g_i(x) \right| \leqslant \lim_{n \to \infty} \left(\frac{2}{3}\right)^{n+1} M_0 = 0$$

d'où $f(x) = \widetilde{f}(x)$, $x \in X$. Ainsi, $\widetilde{f}$ est un prolongement de f tel que, pour $x \in \mathbb{R}^N$,

$$|\widetilde{f}(x)| \leqslant \frac{1}{3} M_0 \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n = M_0$$

d'où $\widetilde{f}$ est un prolongement voulu de f.

Le cas de f non bornée est traité par une transformation appropriée ; posons

$$F(x) = \operatorname{arc\,tg} f(x) , \quad x \in X ,$$

alors

$$F \in C(X) \quad \text{et} \quad |F(x)| < \frac{\pi}{2} , \quad x \in X .$$

Soit $\widetilde{F} \in C(\mathbb{R}^N)$ telle que

$$|\widetilde{F}(x)| \leqslant \frac{\pi}{2} , \quad x \in \mathbb{R}^N .$$

Soit

$$B = \left\{ x \in \mathbb{R}^N : |\widetilde{F}(x)| = \frac{\pi}{2} \right\} ,$$

alors B est un ensemble fermé disjoint de X. Selon le lemme de Urysohn (§ 1.9.1), on a une fonction $\theta \in C(\mathbb{R}^N)$ telle que

$$\theta(x) = 0 \text{ si } x \in B , \quad \theta(x) = 1 \text{ si } x \in X , \quad |\theta(x)| \leqslant 1 , \quad x \in \mathbb{R}^N .$$

Alors la fonction $\widetilde{f}$, définie par

$$\widetilde{f}(x) = \operatorname{tg}\{\theta(x)\widetilde{F}(x)\} , \quad x \in \mathbb{R}^N ,$$

donne un prolongement continu de f. ♦

1.9.4 Fonctions convexes

Soit A une partie convexe de $\mathbb{R}^N$; rappelons que ceci veut dire que si x, y sont dans A, alors le segment $[x, y]$ est dans A, où

$$[x, y] = \{\alpha x + (1-\alpha)y : 0 \leqslant \alpha \leqslant 1\}.$$

On dit que $[x, y]$ est formé des ***combinaisons convexes*** de x et y ; plus généralement, si $x_1, \ldots, x_n$ sont n points de $\mathbb{R}^N$, une ***combinaison convexe*** de $x_1, \ldots, x_n$ est donnée par

$$\alpha_1 x_1 + \cdots + \alpha_n x_n$$

où les $\alpha_i \geqslant 0$ avec $\alpha_1 + \cdots + \alpha_n = 1$.

Notons que *si A est convexe et $x_1, \ldots, x_n$ sont dans A, alors toute combinaison convexe de $x_1, \ldots, x_n$ est dans A.* En effet, pour $n = 2$, c'est la définition de la convexité de A ; en supposant l'affirmation vérifiée pour $n = m$ (hypothèse de récursion), démontrons-la pour $n = m+1$: si $x_1, \ldots, x_{m+1}$ sont dans A et

$$x = \alpha_1 x_1 + \cdots + \alpha_{m+1} x_{m+1}$$

est une combinaison convexe de $x_1, \ldots, x_{m+1}$, on a

$$x = (\alpha_1 + \cdots + \alpha_m)\left\{\frac{\alpha_1}{\alpha_1 + \cdots + \alpha_m} x_1 + \cdots + \frac{\alpha_m}{\alpha_1 + \cdots + \alpha_m} x_m\right\} + \alpha_{m+1} x_{m+1}$$

(en supposant que $\alpha_1 + \cdots + \alpha_m = 1 - \alpha_{m+1} > 0$, le cas $\alpha_{m+1} = 1$ étant trivial) ce qui démontre que

$$x = \lambda z + (1-\lambda) x_{m+1}, \quad \lambda = \alpha_1 + \cdots + \alpha_m,$$

où

$$z = \frac{\alpha_1}{\alpha_1 + \cdots + \alpha_m} x_1 + \cdots + \frac{\alpha_m}{\alpha_1 + \cdots + \alpha_m} x_m.$$

On en déduit d'abord que $z \in A$ (selon l'hypothèse de récurrence) et ensuite $x \in A$ (par convexité de A).

Si p est une norme dans $\mathbb{R}^N$, les ensembles

$$A_1 = \{x : p(x-a) < r\}, \quad A_2 = \{x : p(x-a) \leqslant r\}$$

sont convexes. En effet, si x, y sont dans A_i et u est une combinaison convexe de x, y,

$$u = \alpha x + (1-\alpha)y,$$

alors

$$\begin{aligned} p(u-a) &= p\{\alpha(x-a) + (1-\alpha)(y-a)\} \\ &\leqslant \alpha p(x-a) + (1-\alpha)p(y-a) ; \end{aligned}$$

d'où $u \in A_i$. En particulier, les ensembles

$$B(a;r), \quad \overline{B}(a;r), \quad P(a;r), \quad \overline{P}(a;r)$$

sont convexes ; ils sont définis par les normes $\|\cdot\|$ et $|\cdot|$.

Pour l'ensemble (cube fermé de centre a)

$$\overline{P}(a;r) = \{x : |x-a| \leqslant r\},$$

on pose

$$S = \{x : x = a + r(\varepsilon_1, \ldots, \varepsilon_N),\ \varepsilon_i = \pm 1,\ i = 1, \ldots, N\}.$$

S contient 2^N points, appelés les ***points extrémaux*** du cube fermé $\overline{P}(a;r)$. *Chaque point de $\overline{P}(a;r)$ est une combinaison convexe des points de S.* Pour établir ceci, il suffit de considérer le cas $a = 0$, $r = 1$ et procéder par récurrence ; pour $N = 1$, c'est évident ; aussi, un point du type

$$(\varepsilon_1, \ldots, \varepsilon_{N-1}, t), \quad \varepsilon_i = \pm 1,\ |t| \leqslant 1$$

est une combinaison convexe des deux points

$$(\varepsilon_1, \ldots, \varepsilon_{N-1}, -1), \quad (\varepsilon_1, \ldots, \varepsilon_{N-1}, 1),$$

d'où on peut compléter aisément la preuve par récurrence.

Une fonction réelle f dans un ensemble convexe A dans $\mathbb{R}^N$ s'appelle ***convexe*** si

$$f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha)f(y), \quad x, y \in A,\ 0 \leqslant \alpha \leqslant 1.$$

On démontre que *si f est convexe dans A, alors*

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leqslant \sum_{i=1}^{n} \alpha_i f(x_i)$$

dès que $x_1, \ldots, x_n$ sont dans A et $\alpha_1, \ldots, \alpha_n$ sont $\geqslant 0$ de somme 1. La démonstration de cette dernière est par récurrence sur n ; si $n = 2$, c'est la définition de la convexité de f ; en supposant l'inégalité vraie pour $n = m$, on l'établit pour $n = m + 1$ en écrivant (comme déjà fait ci-dessus, avec $\alpha_{m+1} < 1$)

$$x = \sum_{i=1}^{m+1} \alpha_i x_i = \lambda z + (1-\lambda) x_{m+1}$$

avec z, λ définis comme auparavant. Alors,

$$\begin{aligned} f(x) &= \lambda f(z) + (1-\lambda) f(x_{m+1}) \\ &\leqslant \lambda \left\{ \frac{\alpha_1}{\alpha_1 + \cdots + \alpha_m} f(x_1) + \cdots + \frac{\alpha_m}{\alpha_1 + \cdots + \alpha_m} f(x_m) \right\} \\ &\quad + \alpha_{m+1} f(x_{m+1}) \\ &= \alpha_1 f(x_1) + \cdots + \alpha_m f(x_m) + \alpha_{m+1} f(x_{m+1}). \end{aligned}$$

avec $\lambda = \alpha_1 + \cdots + \alpha_m$, $1 - \lambda = \alpha_{m+1}$.

Nous démontrons maintenant la proposition principale de ce paragraphe.

Si f est convexe dans une partie convexe, ouverte A de $\mathbb{R}^N$ alors f est continue.

DÉMONSTRATION

Si $a \in A$ et K est un cube fermé du type $\overline{P}(a,r)$ contenu dans A, démontrons d'abord que f est borné supérieurement dans K. En effet, si

$$M = \max_{s \in S} f(s)$$

où S est l'ensemble fini des points extrémaux de K, alors

$$f(x) \leqslant M\,, \quad x \in K\,,$$

car chaque $x \in K$ est une combinaison convexe de $s \in S$:

$$x = \sum_{s \in S} \alpha(s)s\,, \quad \alpha(s) \geqslant 0\,, \sum_{s} \alpha(s) = 1\,;$$

d'où

$$f(x) \leqslant \sum_{s} \alpha(s) f(s) \leqslant M\,.$$

Démontrons maintenant la continuité de f à un point quelconque $a \in A$. En passant à la fonction convexe g définie par

$$g(x) = f(x+a) - f(a)\,, \quad x \in A - a = \{b - a : b \in A\}\,,$$

on se ramène à la continuité de g au point 0 avec $g(0) = 0$ où g est définie dans l'ensemble convexe, ouvert $B = A - a$, $0 \in B$. Choisissons un $r > 0$ tel que

$$\overline{P}(0;r) \subset B\,;$$

comme g est borné supérieurement dans $\overline{P}(0,r)$, on a une constante $M > 0$ telle que

$$g(x) \leqslant M\,, \quad \text{si } |x| \leqslant r\,.$$

Si $|x| < \varepsilon r$, $0 < \varepsilon < 1$, écrivons

$$x = \varepsilon \cdot \left(\frac{1}{\varepsilon} x\right) + (1-\varepsilon) \cdot 0\,;$$

on aura, par la convexité de g,

$$g(x) \leqslant \varepsilon g\left(\frac{1}{\varepsilon} x\right) \leqslant \varepsilon M$$

car $\left|(1/\varepsilon)x\right| < r$; en écrivant

$$0 = \frac{1}{1+\varepsilon}\, x + \frac{\varepsilon}{1+\varepsilon}\left(-\frac{1}{\varepsilon} x\right)\,,$$

on a, de même, pour $|x| < \varepsilon r$,

$$0 = g(0) \leqslant \frac{1}{1+\varepsilon}\, g(x) + \frac{\varepsilon}{1+\varepsilon}\, g\left(-\frac{1}{\varepsilon}\, x\right) \leqslant \frac{1}{1+\varepsilon}\, g(x) + \frac{\varepsilon}{1+\varepsilon} M$$

ce qui donne

$$g(x) \geqslant -\varepsilon M\ ;$$

ainsi, au total,

$$\big|g(x)\big| \leqslant \varepsilon M \quad \text{si } |x| < \varepsilon r\,,\ 0 < \varepsilon < 1\,,$$

ce qui montre que g est continue à 0. ♦

Une fonction $f : A \to \mathbb{R}$ (A partie convexe de $\mathbb{R}^N$) s'appelle ***concave*** si $-f$ est convexe, c'est-à-dire

$$f\big(\alpha x + (1-\alpha)y\big) \geqslant \alpha f(x) + (1-\alpha) f(y)\,, \quad x, y \in A\,,\ 0 \leqslant \alpha \leqslant 1\,.$$

Les propriétés des fonctions concaves se déduisent immédiatement de celles des fonctions convexes.

Déjà, pour $N = 1$, une fonction convexe n'est pas nécessairement continue dans un ensemble convexe fermé. Ainsi,

$$f(x) = \begin{cases} 0 & \text{si } 0 < x < 1 \\ 1 & \text{si } x = 0 \text{ ou } 1 \end{cases}$$

est convexe mais discontinue.

Pour $N = 1$, on a le critère de convexité élémentaire bien connu : si f est deux fois dérivables dans un intervalle de $\mathbb{R}$ alors f est convexe si et seulement si $f'' \geqslant 0$. Ce critère sera généralisé au cas $N > 1$ dans la suite.

1.9.5 Un théorème du point fixe

La proposition suivante est très utile lorsqu'on la place dans le cadre d'un espace métrique complet X ; elle est utilisée pour démontrer une grande variété de théorèmes d'existence et d'unicité des solutions d'équations différentielles et intégrales. La démonstration, d'ailleurs, est la même dans le cadre général que dans le cas d'un ensemble fermé de $\mathbb{R}^N$ qui est présenté ici. *Soit X un ensemble fermé non vide de $\mathbb{R}^N$ et $f : X \to X$ une application telle que, pour une constante k, $0 \leqslant k < 1$,*

$$\big\|f(x) - f(y)\big\| \leqslant k\|x - y\|\,, \quad x, y \in X\,.$$

(Une telle application s'appelle une ***contraction*** de constante k.) *Alors il existe un unique point $p \in X$ tel que $f(p) = p$* (p s'appelle un ***point fixe*** pour f).

DÉMONSTRATION

L'unicité de p est immédiate (même si X n'est pas fermé), car, si p' est un autre point fixe pour f, on aura

$$\big\|p - p'\big\| = \big\|f(p) - f(p')\big\| \leqslant k\big\|p - p'\big\|$$

ce qui donne $\big\|p - p'\big\| = 0$ à cause du fait que $0 < k < 1$.

Pour établir l'existence de p, on utilise la méthode dite des ***approximations successives***. On prend un point quelconque x_0 de X et on pose

$$x_1 = f(x_0)\,,\ x_2 = f(x_1)\,,\ \ldots\,,\ x_n = f(x_{n-1})\,,\quad n \geqslant 1\,.$$

Démontrons que $\{x_n\}$ est une suite de Cauchy dans X ; en effet, on vérifie immédiatement que

$$\|x_{n+1} - x_n\| = \|f(x_n) - f(x_{n-1}\| \leqslant k\|x_n - x_{n-1}\|\ ;$$

d'où, par récurrence

$$\|x_{n+1} - x_n\| \leqslant k^n\|x_1 - x_0\|\,,\quad n \geqslant 1\,.$$

Donc, si $n > m$,

$$\begin{aligned}\|x_n - x_m\| &\leqslant \|x_n - x_{n-1}\| + \cdots + \|x_{m+1} - x_m\| \\ &\leqslant \left(k^{n-1} + \cdots + k^m\right)\|x_1 - x_0\| \to 0 \text{ lorsque } m,\, n \to \infty\,,\end{aligned}$$

car $\sum_n k^n < \infty$.

Ainsi, il existe $p \in \mathbb{R}^N$ tel que

$$\lim_{n\to\infty} x_n = p\,.$$

X étant fermé et $\{x_n\}$ étant une suite convergente dans X, on a que $p \in X$; comme f est évidemment continue, on aura

$$f(p) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = p\,. \qquad \blacklozenge$$

1.9.6 Exercices

[1] Si $f \in C(\mathbb{R}^N)$ alors $f^{-1}(0)$ est fermé ; si F est fermé dans $\mathbb{R}^N$, il existe $f \in C(\mathbb{R}^N)$, $0 \leqslant f \leqslant 1$, tel que $F = f^{-1}(0)$.

[2] Si $f \in C(\mathbb{R}^N)$ alors $\{x : f(x) > 0\}$ est ouvert ; si V est ouvert dans $\mathbb{R}^N$, il existe $f \in C(\mathbb{R}^N)$, $0 \leqslant f \leqslant 1$, tel que $V = \{x : f(x) > 0\}$.

[3] Soient $A = \{(x,y) : xy = 1\}$, $B = \{(0,y) : y \in \mathbb{R}\}$ dans $\mathbb{R}^2$; vérifier que A, B sont fermés et disjoints, $d(p,A) > 0$ si $p \in B$ mais $d(A,B) = 0$.

[4] Soit $A = \{(x,y) : x + y = 1\}$ dans $\mathbb{R}^2$, calculer $d\big((0,0)\,,\,A\big)$.

[5] Soient $\{f_n\}$, $\{g_n\}$ deux suites de fonctions réelles dans un ensemble X ; vérifier que si $f_n \to f$, $g_n \to g$ uniformément dans X alors $f_n + g_n \to f + g$ uniformément dans X ; si, de plus, pour un M, $|f_n|$ et $|g_n|$ sont $\leqslant M$, alors $f_n g_n \to fg$ uniformément dans X.

6 La suite de fonctions réelles données par

$$f_n(x) = x^n(1-x)\,, \quad 0 \leqslant x \leqslant 1\,,\ n = 1, 2, \ldots\,,$$

converge uniformément vers 0 dans $[0\,,\,1]$. La suite de fonctions réelles données par

$$g(x) = x^n\big(1-x^n\big)\,, \quad 0 \leqslant x \leqslant 1\,,\ n = 1, 2, \ldots\,,$$

converge vers la fonction continue 0 dans $[0\,,\,1]$ mais non uniformément.

7 Soit V une partie de $\mathbb{R}^N$; démontrer que si $f : V \to \mathbb{R}$ est uniformément continue, alors il existe un unique prolongement $\widetilde{f} : \overline{V} \to \mathbb{R}$ de f.

8 Si $f_1, \ldots, f_n$ sont des fonctions convexes dans un ensemble convexe A, vérifier que les fonctions

$$c_1 f_1 + \cdots + c_n f_n \ (c_1, \ldots, c_n \geqslant 0)\,, \quad \max_{1 \leqslant i \leqslant n} f_i\,,$$

sont aussi des fonctions convexes dans A.

9 Soit $F = \overline{B}(a; r)$ dans $\mathbb{R}^N$ et $f : F \to \mathbb{R}^N$ une contraction de constante k telle que

$$\big\|f(a) - a\big\| \leqslant (1-k)r\,.$$

Vérifier que $f(F) \subset F$; en conclure que f possède un unique point fixe dans F.

1.10 EXERCICES

1 Vérifier que si $B_1 = B(0; r_1)$, $B_2 = B(0; r_2)$ dans $\mathbb{R}^N$, alors $B_1 + B_2 = B(0\,;\, r_1 + r_2)$ où $B_1 + B_2$ est l'ensemble formé par les éléments de la forme $x + y$, $x \in B_1$, $y \in B_2$.

2 Vérifier que $\overline{A \cup B} = \overline{A} \cup \overline{B}$, $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$ pour deux ensembles A, B dans $\mathbb{R}^N$; donner un exemple où $\overline{A \cap B} \neq \overline{A} \cap \overline{B}$.

3 Démontrer que si A, A^c sont non vides (dans $\mathbb{R}^N$) alors $\operatorname{Fr} A \neq \emptyset$.

4 Si $A \subset \mathbb{R}^N$, démontrer que $\operatorname{Fr} A = B$ est fermé ; si F est fermé et $\mathring{F} = \emptyset$ déterminer $\operatorname{Fr} F$.

5 Décrire les ensembles de $\mathbb{R}^N$ dont la frontière est : $\emptyset$, un point, un segment.

6 Soit $f : \mathbb{R}^N \to \mathbb{R}$ tel que, pour tout $a \in \mathbb{R}$, les ensembles

$$\big\{x : f(x) > a\big\}\,, \quad \big\{x : f(x) < a\big\}$$

sont ouverts. Que pouvez-vous dire sur f ?

7 Si x, y sont réels, vérifier que

$$\max(x,y) = \frac{1}{2}\left(x+y+|x-y|\right), \quad \min(x,y) = \frac{1}{2}\left(x+y-|x-y|\right).$$

En déduire que si f et g sont deux fonctions réelles continues (uniformément continues) dans $A \subset \mathbb{R}^N$ alors les fonctions M, m définies par $M(x) = \max\{f(x), g(x)\}$, $m(x) = \min\{f(x), g(x)\}$, sont aussi continues (uniformément continues).

8 Soit $f : \mathbb{R} \to \mathbb{R}$ et A le graphe de f :

$$A = \{(x, f(x)) : x \in \mathbb{R}\} \subset \mathbb{R}^2 .$$

Vérifier que $\mathring{A} = \emptyset$; si f est continue alors A est fermé ; si f est localement bornée et A est fermé, alors f est continue mais, en général, A peut être fermé sans que f soit continue (par exemple, $f(x) = 1/x$ si $x \neq 0$, $f(0) = 0$; f s'appelle *localement bornée* si f est bornée dans un voisinage ouvert de chaque point dans son domaine de définition).

9 Soit $f : \mathbb{R}^2 \to \mathbb{R}$; supposons que, pour tout y, $x \mapsto f(x,y)$ est continue et croissante et que, pour tout x, $y \mapsto f(x,y)$ est continue et croissante ; alors f est continue.

10 Soit $u : \mathbb{R}^N \to \mathbb{R}^M$ une application linéaire ; si p est une norme dans $\mathbb{R}^N$ et q une norme dans $\mathbb{R}^M$, démontrer qu'il existe un $C > 0$ tel que

$$q\big(u(x)\big) \leqslant Cp(x), \quad x \in \mathbb{R}^N .$$

11 Soient $u : \mathbb{R}^N \to \mathbb{R}^N$ une application linéaire et p une norme dans $\mathbb{R}^N$; démontrer que u^{-1} existe si et seulement si, pour un $c > 0$, $p\big(u(x)\big) \geqslant cp(x)$, $x \in \mathbb{R}^N$.

12 Soient $S = \{(x,y) : x^2 + y^2 = 1\} \subset \mathbb{R}^2$, $f(x,y) = x/(1-y)$, $y \neq 1$; vérifier que $f : S \setminus \{(0,1)\} \to \mathbb{R}$ est un homéomorphisme en calculant $f^{-1}(t)$, $t \in \mathbb{R}$, explicitement. Faire un dessin pour comprendre le sens géométrique de f.

13 (***Projection stéréographique***) Soit $S^N = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$; posons $f(x) = (1 - x_{N+1})^{-1}(x_1, \ldots, x_N, 0)$, si $x = (x_1, \ldots, x_{N+1})$. Vérifier que $f : S^N \setminus \{(0, 0, \ldots, 0, 1)\} \to \mathbb{R}^N$ est un homéomorphisme en établissant que, si $\xi = (\xi_1, \ldots, \xi_N)$,

$$f^{-1}(\xi) = \left(\|\xi\|^2 + 1\right)^{-1}\left(2\xi_1, \ldots, 2\xi_N, \|\xi\|^2 - 1\right) .$$

14 Soient $C = \{(x,y,z) : x^2 + y^2 = 1\,,\ z \in \mathbb{R}\} \subset \mathbb{R}^3$, $A = \mathbb{R}^2 \setminus \{(0,0)\}$. Définissons $f : A \to C$ par la formule

$$f(x,y) = \left(\frac{\cos\theta}{r}\,,\ \frac{\sin\theta}{r}\,,\ \ln r\right)$$

où (r,θ) sont les coordonnées polaires de $(x,y) \in \mathbb{R}^2$. Vérifier que f est un homéomorphisme de A sur le cylindre C.

15 Soient $A = A_1 \cup A_2 \subset \mathbb{R}^N$, $f : A \to \mathbb{R}$; supposons que les restrictions de f à A_1 et A_2 sont continues. Démontrer que f est continue si A_1, A_2 sont ouverts ou si A_1, A_2 sont fermés mais non en général.

16 Soit A une partie compacte de $\mathbb{R}^N$ et $\{V_\alpha\}_{\alpha\in J}$ un recouvrement ouvert de A. Alors il existe un $\varepsilon > 0$ (appelé ***un nombre de Lebesgue*** pour le recouvrement ouvert donné de A) ayant la propriété suivante : si x, $y \in A$, $\|x-y\| < \varepsilon$, alors il existe un $\alpha \in J$ tel que x, $y \in V_\alpha$. (Procéder *ab contrario*.)

17 Soit A une partie convexe de $\mathbb{R}^N$ et $f : A \to \mathbb{R}$; vérifier que f est convexe si et seulement si l'ensemble

$$A_f = \{(x,t) : x \in A\,,\ f(x) \leqslant t\} \subset \mathbb{R}^{N+1}$$

est convexe.

18 Soit A une partie fermée, non vide de $\mathbb{R}^N$; posons $f(x) = d(x,A)$, $x \in \mathbb{R}^N$. Démontrer que, pour tout $x \in \mathbb{R}^N$, il existe un $a = a(x)$ dans A tel que $f(x) = \|x - a\|$; établir ensuite que si A est convexe, alors f est une fonction convexe et $a = a(x)$ est uniquement déterminée par x. Réciproquement, si f est une fonction convexe, alors A est un ensemble convexe (voir que $A = \{x : f(x) \leqslant 0\}$).

19 Soit $B : \mathbb{R}^M \times \mathbb{R}^N \to \mathbb{R}^P$ une application ***bilinéaire*** (c'est-à-dire $x \mapsto B(x,y)$, $y \mapsto B(x,y)$ sont des applications linéaires). Démontrer que B est continue ; si $|\cdot|_M$, $|\cdot|_N$ et $|\cdot|_P$ sont trois normes quelconques dans les espaces $\mathbb{R}^M$, $\mathbb{R}^N$, $\mathbb{R}^P$ respectivement, alors il existe une constante $c > 0$ telle que

$$|B(x,y)|_P \leqslant c|x|_M|y|_N\,, \quad x \in \mathbb{R}^M\,,\ y \in \mathbb{R}^N\,.$$

1.11 COMMENTAIRES

Comme mentionné dans le texte (sect. 1.1, § 1.3.4), la plupart des discussions de ce chapitre peuvent être généralisées sans peine aux espaces métriques généraux et même aux espaces topologiques généraux. Ces derniers sont les

ensembles munis d'une famille de parties dites *ouvertes* satisfaisant les conditions (a), (b) et (c) de la proposition 1.2.1. La multiplicité d'applications dans tous les domaines des mathématiques a rendu la notion d'espace topologique comme une des notions centrales des mathématiques. Introduite au début du XX^{e} siècle, l'importance de la topologie n'a cessé d'augmenter partout.

Les espaces vectoriels normés, plus généraux que $\mathbb{R}^N$, sont aujourd'hui un cadre courant pour beaucoup de développements analytiques ; les espaces normés complets (appelés *espaces de Banach*, introduits autour de 1930) et leurs généralisations immédiates sont essentiels pour l'analyse moderne. *Les espaces de Hilbert* (ce sont des espaces normés complets dont la norme provient d'un produit scalaire) sont peut-être encore plus importants, surtout pour leur rôle essentiel dans le développement de la mécanique quantique.

Tous les sujets traités dans ce chapitre sont standards ; on les trouvera présentés dans beaucoup d'ouvrages. Les livres de Dieudonné (tome 1) [13] et de Schwartz (Analyse 1) [38] présentent la grande partie de la matière traitée dans ce chapitre de manière plus détaillée. La démonstration de la continuité des fonctions convexes est basée, en partie, sur un raisonnement dans Bourbaki (Espaces vectoriels topologiques, EVT II.20) [4] ; les ouvrages (collectifs – Bourbaki étant le pseudonyme utilisé par un groupe de mathématiciens éminents français) de ce dernier présentent les versions les plus complètes de toutes les questions considérées ici. Avec un tout petit peu plus d'effort, on peut démontrer que toute fonction convexe f (dans une partie convexe, ouverte) est même *localement lipschitzienne*, c'est-à-dire f satisfait une inégalité de type

$$\left|f(x) - f(y)\right| \leqslant M \cdot \|x - y\|$$

pour x, y dans des boules suffisamment petites. Ceci et bien d'autres propriétés des fonctions convexes peuvent être vues dans le livre de Roberts et Varberg [35] écrit à un niveau comparable à celui du présent ouvrage.

CHAPITRE 2

Dérivabilité

La notion principale de ce chapitre est celle de la dérivée f' en un point d'une fonction f de type $\mathbb{R}^N \to \mathbb{R}^M$. Cette dérivée est définie dans la section 2.1 comme une application linéaire appropriée de $\mathbb{R}^N$ dans $\mathbb{R}^M$; du point de vue numérique, elle est donnée par une matrice de type (M, N) obtenue par le calcul de $M \cdot N$ dérivées partielles associées à f. Au-delà de l'importance théorique de cette notion de dérivée, son utilité se révèle aussi dans la simplicité de la formule pour la dérivation des fonctions composées donnée dans la section 2.2. Les sections 2.3 et 2.4 développent quelques applications importantes de la dérivée. La dernière section 2.6 présente quelques considérations intuitives et introduit certains opérateurs différentiels fondamentaux.

2.1 LA DÉRIVÉE

2.1.1 Généralités

Rappelons d'abord la définition de la dérivée d'une fonction réelle f d'une variable réelle ; si f est définie dans un intervalle ouvert contenant le point $a \in \mathbb{R}$, la dérivée de f au point a, notée $f'(a)$, est définie par la formule

$$f'(a) = \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(a+h) - f(a)}{h}$$

pourvu que la limite existe et est finie. Dans certains raisonnements, il est utile de permettre les valeurs ∞ ou $-\infty$ à la limite $f'(a)$ mais nous n'aurons pas besoin de cette généralisation. Notons que f étant définie dans un intervalle ouvert contenant a, $f(a+h)$ est défini pour toutes les valeurs de h suffisamment petites, de sorte que la limite en question peut être étudiée de manière élémentaire. Si $f'(a) = c$ existe, on voit facilement que l'on aura

$$f(a+h) = f(a) + ch + |h|\,\rho(h)\,, \tag{2.1}$$

où $\rho(h) \to 0$ lorsque $h \to 0$; le nombre c est uniquement déterminé par (2.1). Ainsi, nous pouvons formuler la dérivabilité de f au point a en disant que f satisfait (2.1) pour un nombre réel c approprié.

Si nous posons $u(h) = ch$, nous voyons que u définit une application linéaire de $\mathbb{R}$ dans $\mathbb{R}$; par ailleurs, toute application linéaire u de $\mathbb{R}$ dans $\mathbb{R}$ est de cette forme : à chaque u correspond un unique nombre réel c tel que

$$u(h) = ch\,, \quad h \in \mathbb{R}.$$

On peut donc identifier le nombre c avec l'application linéaire u correspondante et dire que f est dérivable au point a s'il existe une application linéaire $u : \mathbb{R} \to \mathbb{R}$ telle que

$$f(a+h) = f(a) + u(h) + |h|\,\rho(h)\,, \quad \lim_{h\to 0} \rho(h) = 0\,. \tag{2.2}$$

Il est naturel alors de définir $f'(a)$ comme égal à l'application linéaire u ; le fait que u est déterminé par un nombre réel c nous permet de considérer $f'(a)$ tantôt comme l'application linéaire

$$h \longmapsto f'(a)h = u(h)\,,$$

tantôt comme le nombre réel $c = f'(a)$.

L'équation (2.2) peut être interprétée aussi comme suit : f est dérivable au point a si et seulement si l'accroissement $f(a+h)-f(a)$ est approximable par $u(h)$, une expression linéaire en h, l'erreur d'approximation étant d'un ordre de grandeur inférieur à $|h|$ lorsque $h \to 0$. L'idée principale de la derivabilité de f au point a est donc contenue dans la possibilité d'approcher la fonction f par une fonction affine (polynôme de degré un),

$$x \longmapsto f(a) + u(x-a)\,,$$

localement autour du point a dans le sens que

$$\big|f(x) - \{f(a) + u(x-a)\}\big| = |x-a|\,\big|\rho(x-a)\big|$$

avec $\rho(x-a) \to 0$ lorsque $x \to a$.

Cette discussion préliminaire montre la bonne voie pour la définition de la notion de dérivabilité dans le cas où f est une fonction dans $\mathbb{R}^N$ à valeurs dans $\mathbb{R}^M$.

Soit $f : \Omega \to \mathbb{R}^M$, Ω une partie ouverte de $\mathbb{R}^N$. On dit que f est ***dérivable*** (ou ***différentiable***) au point a de Ω, s'il existe une application linéaire $u : \mathbb{R}^N \to \mathbb{R}^M$ telle que

$$f(a+h) = f(a) + u(h) + \|h\|\,\rho(h) \tag{2.3}$$

où $\big\|\rho(h)\big\| \to 0$ lorsque $\|h\| \to 0$. On écrit $f'(a)$ (ou $\mathrm{D}f(a)$) pour l'application linéaire u et on l'appelle la ***dérivée*** (ou ***différentielle***) de f au point a ; on verra immédiatement dans la suite (prop. 2.1.2 et 2.1.4) que l'application linéaire u est uniquement déterminée dès qu'elle existe. Notons aussi que dans (2.3) nous pouvons remplacer le terme $\|h\|\rho(h)$ par $|h|\widetilde{\rho}(h)$ (avec $\big|\widetilde{\rho}(h)\big| \to 0$

lorsque $|h| \to 0$ où $|h|$ est la norme maximum) sans changer la portée de la définition ; nous utilisons cette définition alternative équivalente si cela est plus commode dans un raisonnement.

Soit $\nu \in \mathbb{R}^N$, $\nu \neq 0$; on pose

$$\mathrm{D}_\nu f(a) = \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(a + t\nu) - f(a)}{t} \tag{2.4}$$

si la limite existe. Si $\|\nu\| = 1$, on dit que $\mathrm{D}_\nu f(a)$ est la ***dérivée directionnelle*** de f ***dans la direction*** ν au point a. Si $\nu = e_j$, le j-ième vecteur de la base canonique de $\mathbb{R}^N$, on écrit

$$\mathrm{D}_j f(a) \quad \text{ou} \quad \frac{\partial f}{\partial x_j}(a)$$

au lieu de $\mathrm{D}_{e_j} f(a)$. L'expression $\mathrm{D}_j f(a)$ s'appelle la ***dérivée partielle*** de f ***par rapport à*** x_j au point a. Noter que $\mathrm{D}_j f(a)$, $\mathrm{D}_\nu f(a)$ sont des éléments de $\mathbb{R}^M$.

Si $M = 1$ dans la discussion précédente (c'est-à-dire $f : \Omega \to \mathbb{R}$, Ω une partie ouverte de $\mathbb{R}^N$) et les $\mathrm{D}_j f(a)$, $j = 1, \ldots, N$, existent le vecteur de $\mathbb{R}^N$ donné par

$$\big(\mathrm{D}_1 f(a), \ldots, \mathrm{D}_N f(a)\big)$$

s'appelle le ***gradient*** de f au point a ; on le note aussi par

$$\operatorname{grad} f(a) \quad \text{ou} \quad \nabla f(a)$$

(prononcé : *nabla* $f(a)$). Dans le cas où f est une fonction de deux variables réelles ($N = 2$), une direction ν est souvent fixée par la donnée d'un angle θ entre 0 et 2π (ou entre 0 et π) et en posant

$$\nu = (\cos\theta, \sin\theta) ;$$

dans ce cas, $\mathrm{D}_\nu f(a)$ est souvent notée comme $\mathrm{D}_\theta f(a)$.

Nous dirons que $f : \Omega \to \mathbb{R}^N$ est ***dérivable*** (ou ***différentiable***) dans Ω si f est dérivable en chaque point a de Ω.

2.1.2 Proposition

Soit $f : \Omega \to \mathbb{R}$, Ω une partie ouverte de $\mathbb{R}^N$.

(i) Si f est dérivable au point $a \in \Omega$, f est continue en ce point.

(ii) Si f est dérivable au point $a \in \Omega$, alors les dérivées partielles $\mathrm{D}_j f(a)$, $j = 1, \ldots, N$, existent ; l'application linéaire $f'(a) : \mathbb{R}^N \to \mathbb{R}$ s'identifie avec le vecteur $\nabla f(a)$ via la formule

$$f'(a)(h) = \nabla f(a) \cdot h = \sum_{j=1}^{N} \mathrm{D}_j f(a) h_j , \quad h = (h_1, \ldots, h_N) \in \mathbb{R} ;$$

aussi, $D_\nu f(a)$ *existent pour tout* $\nu \neq 0$ *dans* $\mathbb{R}^N$ *et*

$$D_\nu f(a) = \nabla f(a) \cdot \nu = \sum_{j=1}^{N} \nu_j D_j f(a) = f'(a)(\nu) \tag{2.5}$$

si $\nu = (\nu_1, \ldots, \nu_N)$.

(iii) Si, pour chaque j, $D_j f(x)$ *existe pour tout* x *dans* Ω *et si, de plus,* $D_j f$ *est continue dans* Ω, $j = 1, \ldots, N$, *alors* f *est dérivable dans* Ω.

DÉMONSTRATION

(i) De l'équation (2.3) (pour $M = 1$), on voit que, si f est dérivable au point a,

$$\lim_{h \to 0} f(a + h) = f(a)$$

car $u(h) \to u(0) = 0$ si $h \to 0$; donc f est continue au point a. Il faut se rappeler ici qu'une application linéaire

$$u : \mathbb{R}^N \longrightarrow \mathbb{R}$$

est donnée par

$$u(h) = c \cdot h = \sum_{j=1}^{N} c_j h_j, \quad h = (h_1, \ldots, h_N)$$

où $c = (c_1, \ldots, c_N)$ est un vecteur fixe de $\mathbb{R}^N$.

(ii) En posant $h = t\nu$ dans (2.3), $\nu \neq 0$ dans $\mathbb{R}^N$, t réel, on obtient (en utilisant $u(t\nu) = tu(\nu)$) :

$$f(a + t\nu) - f(a) = tu(\nu) + |t| \, \|\nu\| \, \rho(t\nu) \, ;$$

d'où on déduit immédiatement l'existence de $D_\nu f(a)$ et l'égalité

$$D_\nu f(a) = \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(a + t\nu) - f(a)}{t} = u(\nu) = f'(a)(\nu) \, .$$

En prenant $\nu = e_j$, on obtient l'existence de $D_j f(a)$; si

$$\nu = (\nu_1, \ldots, \nu_N) = \sum_{j=1}^{N} \nu_j e_j$$

on a, par linéarité de $f'(a)$,

$$f'(a)(\nu) = \sum_{j=1}^{N} \nu_j f'(a)(e_j) = \sum_{j=1}^{N} \nu_j D_j f(a) = \nabla f(a) \cdot \nu$$

ce qui donne (2.5).

(iii) Pour simplifier l'écriture, nous présentons la preuve dans le cas $N = 2$. Si x, $x + h$ sont dans Ω,

$$x = (x_1, x_2)\,, \quad h = (h_1, h_2)\,, \quad x + h = (x_1 + h_1\,,\, x_2 + h_2)\,,$$

écrivons

$$\begin{aligned} f(x + h) = f(x) &+ \{f(x_1 + h_1\,,\, x_2 + h_2) - f(x_1 + h_1\,,\, x_2)\} \\ &+ \{f(x_1 + h_1\,,\, x_2) - f(x_1, x_2)\}\,; \end{aligned}$$

d'où, en utilisant le théorème des accroissements finis pour les fonctions réelles d'une variable réelle, on obtient

$$f(x + h) = f(x) + h_2\, \mathrm{D}_2 f(x_1 + h_1\,,\, x_2 + \theta_2 h_2) + h_1\, \mathrm{D}_1 f(x_1 + \theta_1 h_1\,,\, x_2)\,,$$

θ_1, θ_2 étant des nombres appropriés dans $]\,0\,,\,1[$. On en déduit que

$$\begin{aligned} f(x + h) &= f(x) + h_1\{\mathrm{D}_1 f(x) + \rho_1(h)\} + h_2\{\mathrm{D}_2 f(x) + \rho_2(h)\} \\ &= f(x) + \nabla f(x) \cdot h + R(h) \end{aligned}$$

où

$$R(h) = h_1 \rho_1(h) + h_2 \rho_2(h)\,.$$

Notons que $\rho_i(h) \to 0$ lorsque $h \to 0$ à cause de la continuité de $\mathrm{D}_i f$, $i = 1, 2$; mais

$$|R(h)| \leqslant |h| \max\big(|\rho_1(h)|\,,\, |\rho_2(h)|\big)\,, \quad |h| = \max\big((|h_1|\,,\, |h_2|)\big)\,,$$

ce qui montre que

$$R(h) = |h|\, \rho(h)\,, \quad \rho(h) \to 0 \ (h \to 0)$$

et établit la dérivabilité de f au point x. ♦

2.1.3 Remarques

La démonstration précédente montre que si $f : \Omega \to \mathbb{R}$ est dérivable au point a, sa dérivée $f'(a)$ est uniquement déterminée par le gradient $\nabla f(a)$; l'application linéaire

$$f'(a) : \mathbb{R}^N \to \mathbb{R}$$

est donnée par

$$f'(a)(h) = \nabla f(a) \cdot h\,, \quad h \in \mathbb{R}^N\,,$$

où $\nabla f(a)$ est considéré comme un élément de $\mathbb{R}^N$. Pour harmoniser avec le calcul de $f'(a)$ dans le cas le plus général de $f : \Omega \to \mathbb{R}^M$, traité dans la suite, on considérera $\nabla f(a)$ aussi, si nécessaire, comme un vecteur ligne

$$\nabla f(a) = \big[\mathrm{D}_1 f(a), \ldots, \mathrm{D}_N f(a)\big]\,,$$

c'est-à-dire une matrice de type $(1, N)$ avec

$$f'(a)(h) = \nabla f(a) \cdot h = [\mathrm{D}_1 f(a), \ldots, \mathrm{D}_N f(a)] \begin{bmatrix} h_1 \\ \vdots \\ h_N \end{bmatrix}, \quad h = (h_1, \ldots, h_N).$$

A ce sujet, il faut relire nos conventions d'écriture concernant les calculs matriciels dans $\mathbb{R}^N$ données dans le chapitre Conventions § 3 et en particulier ceci : un élément $h = (h_1, \ldots, h_N)$ de $\mathbb{R}^N$ est traité comme une matrice colonne de type $(N, 1)$ dans les calculs matriciels. L'abus de notation indiquée précédemment de considérer $\nabla f(a)$ tantôt comme un vecteur ligne, tantôt comme un élément de $\mathbb{R}^N$ est assez répandu ; il est très commode et il ne crée aucune confusion.

La proposition 2.1.2(iii) donne une condition suffisante facilement applicable pour que f soit dérivable : il suffit que les dérivées partielles $\mathrm{D}_j f$, $j = 1, \ldots, N$, existent et soient continues. Comme le calcul de $\mathrm{D}_j f$ ne pose aucun problème en général, ce critère de dérivabilité est très pratique ; rappelons que pour calculer $\mathrm{D}_j f(a)$, il suffit de considérer f comme une fonction de la variable x_j en fixant les autres variables $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_N$ aux valeurs constantes $a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_N$: posons

$$g(x_j) = f(a_1, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_N), \quad a = (a_1, \ldots, a_N)$$

alors

$$g'(a_j) = \mathrm{D}_j f(a).$$

Une fonction $f : \Omega \to \mathbb{R}$, Ω ouvert dans $\mathbb{R}^N$, telle que $\mathrm{D}_j f$, $j = 1, \ldots, N$, existent et sont continues dans Ω, s'appelle une ***fonction de classe C^1 dans*** Ω et on écrit $f \in C^1(\Omega)$. La partie (iii) de la proposition 2.1.2 s'énonce alors comme suit : *si $f \in C^1(\Omega)$ alors f est dérivable dans Ω.*

Une autre affirmation utile de la proposition 2.1.2 est contenue dans (2.5) : *si f est dérivable*, les dérivées directionnelles $\mathrm{D}_\nu f$ s'obtiennent à partir du gradient de f (donc des $\mathrm{D}_j f$) grâce à (2.5). On peut reformuler (2.5) en disant que l'application

$$\nu \longmapsto \mathrm{D}_\nu f(a), \quad \nu \in \mathbb{R}^N$$

définit une forme linéaire dans $\mathbb{R}^N$, *si f est dérivable* au point a.

Nous allons voir dans les exercices qui suivent que *la simple existence des $\mathrm{D}_j f$ ne garantit pas même la continuité de f* et que, dans l'absence de dérivabilité de f, la formule (2.5) peut être en défaut (Ex. 5-6, § 2.1.6). Un autre exercice (Ex. 7, § 2.1.6) montre que f' peut bien exister sans que les $\mathrm{D}_j f$ soient continues, c'est-à-dire que la proposition 2.1.2(iii) donne seulement une condition suffisante, mais non nécessaire, pour la dérivabilité d'une fonction. Néanmoins, ce sont les fonctions de la classe C^1 qui joueront un rôle principal dans la suite.

On va maintenant généraliser la proposition 2.1.2 au cas d'une fonction vectorielle de type $\mathbb{R}^N \to \mathbb{R}^M$.

2.1.4 Proposition

Soit $f : \Omega \to \mathbb{R}^M$, Ω une partie ouverte de $\mathbb{R}^N$; posons

$$f = (f_1, \ldots, f_M), \quad f_i : \Omega \to \mathbb{R}, \ i = 1, \ldots, M .$$

(i) Si f est dérivable au point $a \in \Omega$, f est continue en ce point.

(ii) La dérivabilité de f au point $a \in \Omega$ équivaut à la dérivabilité des f_i au point a, $i = 1, \ldots, N$.

(iii) Si f est dérivable au point $a \in \Omega$, alors la matrice représentant l'application linéaire

$$f'(a) : \mathbb{R}^N \to \mathbb{R}^M$$

dans les bases canoniques de $\mathbb{R}^N$ et $\mathbb{R}^M$ est

$$\left[\mathrm{D}_j f_i(a)\right], \quad i = 1, \ldots, M, \ j = 1, \ldots, N ;$$

de plus, $\mathrm{D}_\nu f(a)$ existe pour tout $\nu \neq 0$ dans $\mathbb{R}^N$ et

$$\mathrm{D}_\nu f(a) = f'(a)\nu ;$$

En particulier, $f'(a)$ est unique.

DÉMONSTRATION

(i) La continuité de f au point a découle immédiatement de (2.3) ; en effet, de (2.3),

$$\lim_{h \to 0} f(a + h) = f(a)$$

car $u(h) \to u(0) = 0$ lorsque $h \to 0$.

Il faut se rappeler ici qu'une application linéaire

$$u : \mathbb{R}^N \to \mathbb{R}^M$$

est donnée par une unique matrice $[u_{ij}]$:

$$u(h) = \begin{bmatrix} u_{11} & & u_{1N} \\ \vdots & & \vdots \\ u_{M1} & & u_{MN} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_N \end{bmatrix}, \quad h = (h_1, \ldots, h_N) .$$

(ii) Supposons d'abord que f est dérivable au point a ; en comparant la coordonnée numéro i des deux côtés de (2.3) on obtient

$$f_i(a + h) = f_i(a) + u_i(h) + \|h\| \rho_i(h), \quad i = 1, \ldots, M, \tag{2.6}$$

où

$$\rho = (\rho_1, \ldots, \rho_M)$$

$$u_i(h) = \sum_{j=1}^{N} u_{ij} h_j$$

avec $\rho_i(h) \to 0$ lorsque $h \to 0$; on voit de (2.6) que, pour $i = 1, \ldots, M$, f_i est dérivable au point a et que $f_i'(a)$ se confond avec la ligne numéro i de la matrice $[u_{ij}]$ de u d'où (prop. 2.1.2(ii)) :

$$u_{ij} = \mathrm{D}_j f_i(a), \quad j = 1, \ldots, N.$$

D'autre part, si chaque f_i est dérivable au point a, on aura une formule de type (2.6) avec u_i une forme linéaire dans $\mathbb{R}^N$ et $\rho_i(h) \to 0$ lorsque $h \to 0$, $i = 1, \ldots, M$. Mais (2.6) donne (2.3) si l'on pose

$$\rho = (\rho_1, \ldots, \rho_M), \quad u(h) = \big(u_1(h), \ldots, u_M(h)\big)$$

ce qui donne la dérivabilité de f au point a ; notons que la matrice canonique de $f'(a)$ est formée de M lignes dont la ligne numéro i est donnée par $\nabla f_i(a)$, le gradient de f_i au point a.

(iii) Il reste à voir la formule pour $\mathrm{D}_\nu f(a)$. Pour cela, on pose $h = t\nu$, t réel, dans (2.3) ; en utilisant

$$u(t\nu) = tu(\nu)$$

on a immédiatement l'existence de $\mathrm{D}_\nu f(a)$ et la relation

$$\mathrm{D}_\nu f(a) = u(\nu) = f'(a)(\nu). \qquad \blacklozenge$$

2.1.5 Remarques

La matrice de type (M, N)

$$\big[\mathrm{D}_j f_i(a)\big], \quad i = 1, \ldots, M, \; j = 1, \ldots, N$$

qui apparaît dans la proposition 2.1.4 comme la matrice canonique de $f'(a)$ s'appelle la ***matrice jacobienne*** ou la ***matrice fonctionnelle*** de f au point a. Lorsque $M = N$, la matrice jacobienne est une matrice carrée ; son déterminant s'appelle le ***jacobien*** ou le ***déterminant fonctionnel*** de f au point a ; on le note

$$\det f'(a) \quad \text{ou} \quad \frac{\partial(f_1, \ldots, f_N)}{\partial(x_1, \ldots, x_N)}(a) \quad \text{ou} \quad J_f(a).$$

Le jacobien d'une application de type $\mathbb{R}^N \to \mathbb{R}^N$ joue un rôle important dans la formule de transformation des variables dans les intégrales multiples (chap. 5).

Le cas particulier $N = 1$, $M \geqslant 1$ est simple mais important ; il s'agit d'une fonction

$$f : \Omega \to \mathbb{R}^M, \quad f = (f_1, \ldots, f_M)$$

où, en général, on restreint Ω à être un intervalle ouvert dans $\mathbb{R}$. Une telle fonction donne un chemin dans $\mathbb{R}^M$ (une trajectoire physique d'une particule si $M = 3$). Si $a \in \Omega$,

$$f'(a) = \big(f_1'(a), \ldots, f_M'(a)\big) = {}^t[f_1'(a), \ldots, f_M'(a)]$$

est une matrice de type $(M,1)$; $f'(a)$ existe si et seulement si les M fonctions réelles $f_1, \ldots, f_M$ sont dérivables au sens usuel de la dérivation d'une fonction d'une variable réelle. On remarquera que ceci est en concordance avec le fait qu'une application linéaire

$$u : \mathbb{R} \to \mathbb{R}^M$$

est donnée par

$$u(t) = tc, \quad c = (c_1, \ldots, c_M) \text{ fixe dans } \mathbb{R}^M .$$

On utilisera ces considérations du cas $N = 1$, aussi dans la situation ou Ω est un intervalle fermé ; si $\Omega = [a, b]$, $f'(a)$ sera donnée par la dérivée à droite et $f'(b)$ sera donnée par la dérivée à gauche, quand ces dérivées existent. Si pour un point a de l'intervalle Ω, $f'(a) \neq 0$, on interprète ceci en disant que $f'(a)$ donne la direction de la tangente au chemin décrit par f ; la tangente elle-même est la droite dans $\mathbb{R}^M$ passant par $f(a)$, donnée par l'équation paramétrique

$$\xi = f(a) + tf'(a), \quad t \in \mathbb{R}.$$

Une fonction $f : \Omega \to \mathbb{R}^M$, $f = (f_1, \ldots, f_M)$, Ω une partie ouverte de $\mathbb{R}^N$, est de ***classe*** C^1 dans Ω si chaque fonction composante f_i est de classe C^1 dans Ω ; on écrira $f \in C^1(\Omega, \mathbb{R}^M)$ pour indiquer que f est de classe C^1 dans Ω. Ainsi

$$f \in C^1(\Omega, \mathbb{R}^M) \quad \text{si et seulement si } f_i \in C^1(\Omega),\ i = 1, \ldots, M .$$

Des propositions 2.1.2 et 2.1.4 on conclut que *si* $f \in C^1(\Omega, \mathbb{R}^M)$ *alors* f *est dérivable dans* Ω. Dans la plupart des considérations dans la suite c'est la condition d'appartenance à la classe C^1 qui sera généralement utilisée. Notons que si f est de classe C^1, les fonctions J_f (déterminant fonctionnel de f) ainsi que la fonction matricielle $[\mathrm{D}_j f_i]$ (considérée comme une fonction à valeur dans $\mathbb{R}^{MN}$) sont continues.

2.1.6 Exercices

1 Vérifier que les expressions suivantes définissent des fonctions réelles dérivables et calculer leurs dérivées partielles et leurs gradients :

(i) $\arcsin xy$;
(ii) $\cos(x + y + z)$;
(iii) $\sin(x^2 + y^2) + \cos z$;
(iv) x^y, $x > 0$;
(v) $\ln(x + y\cos z)$, $x > y$;
(vi) $\|x\|$ où $x \in \mathbb{R}^N$, $x \neq 0$;
(vii) $\mathrm{e}^{-\|x\|^2}$, $x \in \mathbb{R}^N$;
(viii) $c \cdot x$, $x \in \mathbb{R}^N$ ($c \in \mathbb{R}^N$ fixe).

2 Calculer $f'(t)$ pour les fonctions f suivantes :

(i) $(r\cos t, r\sin t)$, $0 \leqslant t \leqslant 2\pi$ $(r > 0)$;

(ii) $(r\cos \alpha t, r\sin \alpha t, \beta t)$, $t \in \mathbb{R}$ $(r > 0,\ \alpha \in \mathbb{R},\ \beta \in \mathbb{R}^*)$.

Vérifier que : $f'(t) \neq 0$ pour tout t ; pour (i) $f'(t) \cdot f(t) = 0$ pour tout t ; dans (i) et (ii), $\|f'(t)\|$ reste constante.

3

(i) Si $f(x,y) = x \operatorname{arc\,tg}(x+y)$ calculer $\mathrm{D}_2 f(0,1)$;

(ii) si $f(x,y,z) = \exp\bigl(\sin(x^2+yz)\bigr)$ calculer $\mathrm{D}_1 f(0,1,2)$.

4 Calculer f' comme une matrice appropriée pour les fonctions f suivantes :

(i) $f(x,y) = (xy\,,\ x+y\,,\ x-y)$, $(x,y) \in \mathbb{R}^2$;

(ii) $f(x,y,z) = (xyz\,,\ x+y+z)$, $(x,y,z) \in \mathbb{R}^3$;

(iii) $f(x) = Ax + b$, $x \in \mathbb{R}^N$, (A matrice (M,N), $b \in \mathbb{R}^M$) ;

(iv) $f(x) = \bigl(\|x\|^2, x\bigr) \in \mathbb{R}^{N+1}$, $x \in \mathbb{R}^N$.

5 Soit $f(x,y) = xy^2/(x^2+y^4)$ si $(x,y) \neq (0,0)$, $f(0,0) = 0$, où $(x,y) \in \mathbb{R}^2$. Vérifier que :

(i) f est bornée dans $\mathbb{R}^2$;

(ii) f est discontinue à $(0,0)$;

(iii) $\mathrm{D}_\nu f(0,0)$ existe pour toute direction ν,

$$\nabla f(0,0) = (0,0)\,, \quad \mathrm{D}_\nu f(0,0) \neq \nabla f(0,0) \cdot \nu\ ;$$

(iv) f est de classe C^1, donc dérivable partout sauf au point $(0,0)$.

(Pour (i), voir que $\sup_{x,y}\bigl|f(x,y)\bigr| = \sup_t |t|/(1+t^2) = 1/2$; pour (ii) calculer $f(x,mx)$ $(x \to 0)$, $f(y^2,y)$.)

6 Soit $f(x,y) = yx^3/(x^4+y^2)$ si $(x,y) \neq (0,0)$, $f(0,0) = 0$, où $(x,y) \in \mathbb{R}^2$. Vérifier que :

(i) f est continue partout ;

(ii) $\mathrm{D}_\nu f(0,0) = 0$ pour tout $\nu \neq (0,0)$;

(iii) f n'est pas dérivable au point $(0,0)$;

(iv) f est de classe C^1, donc dérivable, partout sauf au point $(0,0)$.

(Pour (i), voir que $\bigl|f(x,y)\bigr| \leqslant |x|/2$ par inégalité de moyenne arithmétique et géométrique ; pour (iii), il suffit de voir que $f(x,x^2) + \sqrt{x^2+x^4} \not\to 0$ si $x \to 0$.)

7 Soit $f(x,y) = \varphi(r)$, $r = \sqrt{x^2+y^2}$, $(x,y) \in \mathbb{R}^2$, où $\varphi : \mathbb{R} \to \mathbb{R}$ est une fonction dérivable telle que $\varphi(0) = 0 = \varphi'(0)$ mais φ' est discontinue au point 0 (par exemple, $\varphi(t) = t^2 \sin(1/t)$ si $t \neq 0$, $\varphi(0) = 0$). Vérifier que :

(i) f est de classe C^1 partout sauf au point $(0,0)$; $\mathrm{D}_1 f$, $\mathrm{D}_2 f$ existent partout mais ne sont pas continues au point $(0,0)$;

(ii) f est dérivable partout.

8 Soit $f : \Omega \to \mathbb{R}$, Ω ouvert dans $\mathbb{R}^N$; si f est dérivable au point a de Ω, alors

$$\big\|\nabla f(a)\big\| = \sup\big\{\mathrm{D}_\nu f(a) : \|\nu\| = 1\big\}\,.$$

(Ainsi, $\nabla f(a) \neq 0$ donne la direction dans laquelle le taux de variation de f au point a est maximale.)

9 Soient f, g deux applications de Ω dans $\mathbb{R}^M$, Ω étant une partie ouverte de $\mathbb{R}^N$; si f, g sont dérivables au point $a \in \Omega$, démontrer que $f+g$ est dérivable au point a et

$$(f+g)'(a) = f'(a) + g'(a)\,.$$

10 Soient $f : \Omega \to \mathbb{R}^M$, $g : \Omega \to \mathbb{R}^M$, $\varphi(x) = f(x) \cdot g(x)$, $x \in \Omega$, où Ω est une partie ouverte de $\mathbb{R}^N$. Supposons que f, g sont dérivables au point $a \in \Omega$; démontrer que $\varphi : \Omega \to \mathbb{R}$ est dérivable au point a et

$$\mathrm{D}_i\varphi(a) = \mathrm{D}_i f(a) \cdot g(a) + f(a) \cdot \mathrm{D}_i g(a)\,, \quad i = 1, \ldots, N\,,$$

$$\nabla\varphi(a) \cdot h = f'(a)h \cdot g(a) + f(a) \cdot g'(a)h\,, \quad h \in \mathbb{R}^N\ ;$$

en prenant $N = 1$, on aura $\varphi'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a)$ d'où, pour le cas particulier $\varphi(x) = \big\|f(x)\big\|^2$, $\varphi'(a) = 2f(a) \cdot f'(a)$.

11 Soient $f : \Omega \to \mathbb{R}^3$, $g : \Omega \to \mathbb{R}^3$, $\varphi(x) = f(x) \times g(x)$, $x \in \Omega$, où Ω est une partie ouverte de $\mathbb{R}^N$. Supposons que f, g sont dérivables au points $a \in \Omega$; démontrer que $\varphi : \Omega \to \mathbb{R}^3$ est dérivable au point a et

$$\varphi'(a)h = \big\{f'(a)h \times g(a)\big\} + \big\{f(a) \times g'(a)h\big\}\,, \quad h \in \mathbb{R}^N\ ;$$

en prenant $N = 1$, on aura $\varphi'(a) = f'(a) \times g(a) + f(a) \times g'(a)$.

12 Soit $f(x) = Ax \cdot x$, $x \in \mathbb{R}^N$ où $A = \big[a_{ij}\big]$ est une matrice (N,N) ($f(x)$ s'appelle une *forme quadratique* en x). Démontrer, à partir de la définition de f', que $f : \mathbb{R}^N \to \mathbb{R}$ est dérivable et que $f'(x) = (A + {}^tA)x$, ce qui donne $f'(x) = 2Ax$ si A est symétrique. (Noter que, même si A n'est pas symétrique, $f(x) = Bx \cdot x$ où $B = (1/2)(A + {}^tA)$.) Vérifier que

$$f(x) = \sum_{i,j} a_{ij} x_i x_j$$

d'où obtenir la formule pour $f'(x) = \nabla f(x)$ directement.

2.2 DÉRIVATION DES FONCTIONS COMPOSÉES

Pour les fonctions de type $\mathbb{R} \to \mathbb{R}$, on sait que les formules de style

$$(f+g)' = f' + g' , \quad (fg)' = f'g + fg' , \quad (g \circ f)' = (g' \circ f)f'$$

jouent un rôle essentiel pour toutes sortes de calculs. Nous allons voir que les formules analogues sont valables pour les fonctions de type $\mathbb{R}^N \to \mathbb{R}^M$ (Ex. 9-11, § 2.1.6). En fait, nous nous apercevrons qu'au fond toutes ces formules sont des cas particuliers d'une seule formule, celle de la *dérivation des fonctions composées*, qui s'écrit, formellement, exactement de la même manière que dans le cas d'une variable :

$$(g \circ f)' = (g' \circ f)f' .$$

Il suffit d'interpréter cette formule comme la composition (multiplication) des applications linéaires (matrices) appropriées. C'est la ***formule la plus importante*** de la théorie de dérivation des fonctions de plusieurs variables. Une remarque élémentaire s'impose : pour éviter les écritures insensées, il faut toujours vérifier d'abord qu'une expression de style $g\big(f(x)\big)$ est bien définie ; ceci veut dire que $f(x)$ doit être bien dans le domaine de définition de g. Une expression de type $\ln(-x^2)$ n'a aucun sens dans le domaine de nombres réels, même si l'on peut « calculer sa dérivée » comme $2/x$ — un calcul dépourvu de tout intérêt.

2.2.1 Proposition (*formule de dérivation des fonctions composées*)

Soient

$$f : \Omega \to \mathbb{R}^M , \quad g : V \to \mathbb{R}^P ,$$

où Ω est une partie ouverte de $\mathbb{R}^N$, V une partie ouverte de $\mathbb{R}^M$ et $f(\Omega) \subset V$. Posons

$$\varphi = g \circ f , \quad \varphi : \Omega \to \mathbb{R}^P .$$

Si f est dérivable au point a de Ω et g est dérivable au point $f(a)$ de V, alors φ est dérivable au point a et

$$\varphi'(a) = g'\big(f(a)\big) f'(a) . \tag{2.7}$$

DÉMONSTRATION

Pour $h \in \mathbb{R}^N$, $s \in \mathbb{R}^M$ appropriés, on a, en posant $b = f(a)$,

$$f(a+h) = f(a) + f'(a)h + |h|\,\rho(h) ,$$
$$g(b+s) = g(b) + g'(b)s + |s|\,\sigma(s) ,$$

où $\rho(h) \to 0$ lorsque $h \to 0$ et $\sigma(s) \to 0$ lorsque $s \to 0$. Donc

$$\begin{aligned}\varphi(a+h) &= g\big(f(a+h)\big) = g(b+s)\\ &= g(b) + g'(b)\big\{f'(a)h + |h|\,\rho(h)\big\} + |s|\,\sigma(s)\\ &= \varphi(a) + g'(b)f'(a)h + |h|\,\tau(h)\end{aligned}$$

où

$$s = f'(a)h + |h|\,\rho(h) \quad \text{et} \quad \tau(h) = g'(b)\rho(h) + \frac{|s|}{|h|}\,\sigma(s)\,.$$

Pour voir que $\varphi'(a)$ existe et pour établir (2.7) il suffit de montrer que $\tau(h) \to 0$ lorsque $h \to 0$. Comme $\rho(h) \to 0$ lorsque $h \to 0$, on a, par la continuité de l'application linéaire $g'(b)$, que

$$g'(b)\rho(h) \longrightarrow 0 \quad \text{lorsque} \quad h \to 0\,;$$

aussi, lorsque $h \to 0$, on a que $s \to 0$ et $\sigma(s) \to 0$; la démonstration sera donc achevée si nous montrons que $|s|/|h|$ reste bornée lorsque $h \to 0$. Or

$$|s| = \big|f'(a)h + |h|\,\rho(h)\big| \leqslant c|h| + |h|\,\big|\rho(h)\big|$$

où $c > 0$ est tel que $\big|f'(a)h\big| \leqslant c|h|$, $h \in \mathbb{R}^N$; de plus, comme $\rho(h) \to 0$ lorsque $h \to 0$, on aura, pour ε, c' appropriés,

$$\big|\rho(h)\big| \leqslant c' \quad \text{si } |h| \leqslant \varepsilon\,;$$

au total,

$$|s| \leqslant |h|(c+c') \quad \text{si } |h| \leqslant \varepsilon$$

ce qui montre que $|s|/|h|$ est borné lorsque $h \to 0$ et complète notre démonstration. ♦

2.2.2 Remarques

On va interpréter (2.7) en introduisant les fonctions composantes de f, g et φ dans la proposition 2.2.1 :

$$f = (f_1, \ldots, f_M)\,, \quad g = (g_1, \ldots, g_P)\,, \quad \varphi = (\varphi_1, \ldots, \varphi_P)$$

$$\varphi_i(x) = g_i\big(f_1(x), \ldots, f_M(x)\big)\,, \quad x = (x_1, \ldots, x_N)\,.$$

La formule (2.7) équivaut à la suivante :

$$\mathrm{D}_j\varphi_i(x) = \sum_{k=1}^{M} \mathrm{D}_k g_i\big(f(a)\big)\,\mathrm{D}_j f_k(a)\,, \quad i = 1, \ldots, P\,,\ j = 1, \ldots, N\,, \tag{2.8}$$

ce qui correspond à l'interprétation matricielle de (2.7) :

$$\big[\mathrm{D}_j\varphi_i(a)\big]_{(P,N)} = \big[\mathrm{D}_j g_i\big(f(a)\big)\big]_{(P,M)} \cdot \big[\mathrm{D}_j f_i(a)\big]_{(M,N)}\,.$$

La formule (2.8) s'écrit couramment comme suit :

$$\frac{\partial \varphi_i}{\partial x_j} = \sum_{k=1}^{M} \frac{\partial g_i}{\partial y_k} \frac{\partial y_k}{\partial x_j} \tag{2.9}$$

où

$$\varphi_i(x) = g_i(y_1, \ldots, y_M), \quad y_k = f_k(x_1, \ldots, x_N)$$

et où l'on a supprimé les points d'évaluation, a et $f(a)$, des dérivées partielles concernées. Toute formule équivalente à (2.7), (2.8) ou (2.9) s'appelle la formule pour la dérivation des fonctions composées. Dans sa forme suggestive (mais moins précise) (2.9), elle est aussi appelée « *chain rule* » ou « *Kettenregel* » (la règle de dérivation en chaîne).

Notons le cas particulier de (2.7) ou de (2.8), (2.9) : si

$$\varphi(t) = F\big(\gamma_1(t), \ldots, \gamma_N(t)\big), \quad t \in J$$

où F est une fonction différentiable de N variables réelles et $\gamma = (\gamma_1, \ldots, \gamma_N)$ est une fonction différentiable de la variable réelle t parcourant un intervalle J (telles que $F \circ \gamma$ se compose bien dans la manière décrite dans la proposition 2.2.1, c'est-à-dire F est dérivable dans un ouvert Ω qui contient toutes les valeurs $\gamma(t)$, $t \in J$), alors

$$\varphi'(t) = F'\big(\gamma(t)\big)\gamma'(t) = \nabla F\big(\gamma(t)\big) \cdot \gamma'(t). \tag{2.10}$$

Nous laissons la vérification instructive de (2.10) comme un exercice.

2.2.3 Manipulation des différentielles

Il y a un procédé très répandu, de la « manipulation des différentielles », qui contient la formule pour la dérivation des fonctions composées et qui a une interprétation physique assez simple. La théorie mathématique des formes différentielles, que nous considérerons dans un chapitre ultérieur, peut être considérée comme une justification de la validité du procédé. Nous allons la présenter ici brièvement, sans souci de rigueur, en traitant un cas particulier typique.

Soit g une fonction réelle des variables $y_1, \ldots, y_M$ où chaque y_k est une fonction des variables $x_1, \ldots, x_N$. Ecrivons $y_k = f_k(x_1, \ldots, x_N)$; il s'agit d'étudier

$$g\big(f(x)\big) = g\big(f_1(x), \ldots, f_M(x)\big), \quad x = (x_1, \ldots, x_N).$$

On calcule comme suit :

$$\begin{aligned} \mathrm{d}g &= \sum_{k=1}^{M} \mathrm{D}_k g \, \mathrm{d}y_k \\ &= \sum_{k=1}^{M} \mathrm{D}_k g \left\{ \sum_{j=1}^{N} \mathrm{D}_j f_k \, \mathrm{d}x_j \right\} = \sum_{j=1}^{N} \left\{ \sum_{k=1}^{M} \mathrm{D}_k g \, \mathrm{D}_j f_k \right\} \mathrm{d}x_j . \end{aligned}$$

La première égalité ci-dessus est une forme suggestive pour

$$g(y+s) - g(y) = \sum_{k=1}^{M} \mathrm{D}_k g(y) s_k + \text{reste}$$

et la dernière égalité donne la formule de dérivation de la fonction composée $g \circ f$. On peut considérer la proposition 2.2.1 comme une justification mathématique de la manipulation formelle indiquée.

L'expression

$$\mathrm{d}g = \mathrm{D}_1 g \, \mathrm{d}y_1 + \cdots + \mathrm{D}_M g \, \mathrm{d}y_M$$

s'appelle souvent la *différentielle totale* de g (calculée à un point a si les $\mathrm{D}_i g$ sont évalués en ce point) ; cette expression possède une certaine invariance lorsque l'on remplace les y_k par d'autres variables x_j, dont les y_k sont des fonctions. C'est ce fait qui est exprimé dans le calcul formel précédent ; ainsi, quitte à faire des interprétations justes, on écrit

$$\mathrm{d}g = \mathrm{D}_1 g \, \mathrm{d}y_1 + \cdots + \mathrm{D}_M g \, \mathrm{d}y_M = \mathrm{D}_1 \widetilde{g} \, \mathrm{d}x_1 + \cdots + \mathrm{D}_N \widetilde{g} \, \mathrm{d}x_N$$

où $y = f(x)$ et $\widetilde{g} = g \circ f$.

2.2.4 Exercices

1 Soit $f : \mathbb{R}^3 \to \mathbb{R}^2$ telle que $f(0,0,0) = (-1,1)$ et

$$f'(0,0,0) = \begin{bmatrix} -1 & 3 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Soit $g(x,y) = (2x - 3y - 1\,,\, -x^2 y + 2)$; calculer $(g \circ f)'(0,0,0)$.

2 Soient

$$f(x_1,x_2,x_3,x_4) = -2x_1^2 x_2 + 4x_3 - \frac{5x_1}{x_4}\,, \quad (x_1,x_2,x_3,x_4) \in \mathbb{R}^4$$

$$g(x) = \left(x^2,\, -x+1,\, 0,\, (x+1)^3\right), \quad x \in \mathbb{R}.$$

Calculer $(g \circ f)'(0,1,0,1)$, $(f \circ g)'(0)$ et vérifier que

$$\mathrm{Tr}(g \circ f)'(0,1,0,1) = (f \circ g)'(0)\,.$$

3 Soit $g : \mathbb{R}^2 \to \mathbb{R}^3$ telle que $g(0,0) = (0,1,1)$ et

$$g'(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & -1 \end{bmatrix}.$$

Soit

$$f(x,y,z) = \left(\frac{y}{\sqrt{1-y^2}}\, \mathrm{e}^x + z\,,\, y \arctg(\arctg \sqrt{1-y^2}\,)\,,\, z \tg y \right).$$

Calculer $(f \circ g)'(0,0)$.

4 Soient

$$f(x,y,z) = \Big(\frac{x}{2} + yz^2 - 1\,,\; x+y\Big)$$
$$g(x,y) = \big(y \arcsin x\,,\; -x \operatorname{arc\,tg} y + 1\,,\; \operatorname{arc\,sh}(x-y)\big)$$

où x, y, z sont réels ; calculer $(f \circ g)'(0,0)$.

5 Soient

$$f(x,y) = \Big(-x\,\mathrm{e}^{-x+y^2}\,,\; \sin x - 3xy\,,\; \big(x-y^2\big)^3\Big)$$
$$g(x,y,z) = \Big(\big(x-y+z\big)^2\,,\; x+4xy-z\Big) \qquad x,\, y,\, z \in \mathbb{R}.$$

Calculer $(g \circ f)'(0,0)$, $(f \circ g)'(0,0,0)$.

6 Soit $f : \Omega \to \mathbb{R}$, Ω une partie ouverte de $\mathbb{R}^N$ et $\gamma : J \to \Omega$, J un intervalle de $\mathbb{R}$; supposons que f et γ sont dérivables et que $f\big(\gamma(t)\big)$ est constante pour $t \in J$. Vérifier que $\nabla f\big(\gamma(t)\big) \cdot \gamma'(t) = 0$, $t \in J$.

7 Soient $f(u,v) = (u+v,\, u+v,\, u+v)$, $g(u,v) = \big((u+v),\, (u+v)^2,\, (u+v)^3\big)$, u, v réels. Vérifier que $\operatorname{rang} f'(u,v) = 1 = \operatorname{rang} g'(u,v)$, $(u,v) \in \mathbb{R}^2$.

8 Soient $f(u,v) = (u,\, v,\, u+v)$, $g(u,v) = \big(u+v,\, u-v,\, uv\big)$, u, $v \in \mathbb{R}$; vérifier que $\operatorname{rang} f'(u,v) = 2 = \operatorname{rang} g'(u,v)$, $(u,v) \in \mathbb{R}^2$.

9 Soit $\Omega = \big\{x = (x_1, \ldots, x_N) : x_i > 0\,,\; i = 1, \ldots, N\big\}$; si $f : \Omega \to \mathbb{R}$ on dit que f est *positivement homogène* d'ordre $\alpha \in \mathbb{R}$ si

$$f(tx) = t^\alpha f(x)\,, \quad t > 0\,,\; x \in \Omega\,.$$

Pour x fixe dans Ω, posons $\varphi(t) = f(tx)$, $t > 0$. En supposant f dérivable dans Ω on obtient

$$\varphi'(t) = \nabla f(tx) \cdot x = \alpha t^{\alpha-1} f(x)$$

d'où, pour $t = 1$, $\nabla f(x) \cdot x = \alpha f(x)$, $x \in \Omega$ (*théorème d'Euler*).

10 Soit $\varphi(x) = f\big(g(x)\big)$ où $g : \Omega \to \mathbb{R}^N$, $f : V \to \mathbb{R}$, Ω, V étant les parties ouvertes de $\mathbb{R}^N$ telles que $g(\Omega) \subset V$; on suppose que f, g sont dérivables dans leurs domaines de définition. Alors

$$\nabla\varphi(x) = \nabla f\big(g(x)\big) g'(x)\,, \quad x \in \Omega\,;$$

écrire cette relation en termes matriciels, pour le cas $g(x) = Ax$, A une matrice (N,N). En particulier, si A est une matrice orthogonale, vérifier que

$$\big\|\nabla\varphi(x)\big\| = \big\|\nabla f(Ax)\big\|\,, \quad x \in \Omega\,.$$

11 Soient Ω une partie ouverte de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}^N$, $a \in \Omega$, $f(a) = b$. Supposons que g est une application définie dans un voisinage ouvert de b telle que

$$g\big(f(x)\big) = x$$

pour tout x dans un voisinage ouvert de a. Si f est dérivable au point a et g est dérivable au point b, démontrer que $f'(a)$, $g'(b)$ sont inversibles et

$$g'(b) = \big(f'(a)\big)^{-1}, \quad \det g'(a) = \frac{1}{\det f'(a)}\,.$$

2.3 FORMULE DES ACCROISSEMENTS FINIS

La formule invoquée dans le titre de cette section est généralement présentée, dans le cadre des fonctions réelles f d'une variable réelle, comme suit : on a

$$f(b) - f(a) = (b - a)f'(c)$$

si f est continue dans $[a\,,\,b]$ et dérivable dans $]a\,,\,b[$ où c est un point approprié entre a et b. L'origine de la terminologie habituelle « accroissements finis » s'explique historiquement par le fait que la formule donne une relation concernant les « accroissements finis » $f(b) - f(a)$ en opposition à la relation concernant les « accroissements infinitésimaux »

$$f(x + \mathrm{d}x) - f(x) = f'(x)\,\mathrm{d}x$$

qui n'est que la définition de $f'(x)$. La formule des accroissements finis s'appelle aussi le *théorème* ou la *formule de Lagrange* ou encore le *théorème de la moyenne* (cette dernière appellation étant dominante en anglais, allemand ou italien : *mean value theorem*, *Mittelwertsatz*, *teorema del valor medio*). La terminologie « théorème de la moyenne » s'explique par le fait que la formule concernée est presque équivalente à la formule suivante : on a

$$f(b) - f(a) = \int_a^b f'(t)\,\mathrm{d}t$$

en supposant, par exemple, que f' existe et est continue dans $[a\,,\,b]$; ainsi la valeur moyenne de f' sur $[a\,,\,b]$,

$$\frac{1}{b-a}\int_a^b f'(t)\,\mathrm{d}t$$

d'une part est égale à la pente de la droite entre $\big(a, f(a)\big)$ et $\big(b, f(b)\big)$, soit

$$\frac{f(b) - f(a)}{b - a}\,,$$

d'autre part est égale à $f'(c)$ pour un c approprié dans $]a, b[$, cette dernière conclusion étant une conséquence immédiate de la continuité de f' dans $[a, b]$.

L'utilisation essentielle de la formule des accroissements finis passe généralement par une inégalité qui en est une conséquence immédiate : on a (sous les conditions indiquées ci-dessus) :

$$|f(b) - f(a)| \leqslant K(b - a)$$

si $K = \sup_{a<x<b} |f'(x)|$.

Pour les fonctions de type $\mathbb{R}^N \to \mathbb{R}^M$, c'est cette formulation en forme d'une inégalité qui est la plus importante ; par ailleurs, elle se déduit facilement de la formule des accroissements finis dans le cas élémentaire décrit au début de cette section.

2.3.1 Proposition

Soient Ω un ensemble ouvert de $\mathbb{R}^N$, a, b deux points de Ω ; on suppose que le segment $[a, b]$ est contenu dans Ω.

(i) Si $f : \Omega \to \mathbb{R}$ est une fonction, continue dans $[a, b]$ et dérivable à chaque point de $]a, b[$, alors il existe un point c dans $]a, b[$ tel que

$$f(b) - f(a) = (b - a) \cdot \nabla f(c) ;$$

si

$$K = \sup \left\{ \sum_{j=1}^{N} |\mathrm{D}_j f(x)| : x \in]a, b[\right\}$$

alors

$$|f(b) - f(a)| \leqslant K|b - a| .$$

(ii) Si $f : \Omega \to \mathbb{R}^M$, $f = (f_1, \ldots, f_M)$, est une fonction, continue dans $[a, b]$ et dérivable à chaque point de $]a, b[$, alors

$$|f(b) - f(a)| \leqslant K|b - a|$$

où

$$K = \sup \left\{ \sum_{j=1}^{N} |\mathrm{D}_j f_i(x)| : i = 1, \ldots, M , \ x \in]a, b[\right\} = \sup_{a<x<b} \|f'(x)\| .$$

DÉMONSTRATION

(i) L'idée principale de la preuve est de considérer la fonction réelle d'une variable réelle t dans $[0, 1]$ définie par

$$\varphi(t) = f\big(a + t(b - a)\big) .$$

Sous les hypothèses faites sur f, la fonction φ sera continue dans $[0, 1]$ et dérivable sur $]0, 1[$; de plus

$$\varphi(0) = f(a), \quad \varphi(1) = f(b).$$

En se servant de la formule des accroissements finis déjà invoquée au début de cette section, on a que

$$f(b) - f(a) = \varphi(1) - \varphi(0) = \varphi'(\theta)$$

pour un θ approprié dans $]0, 1[$; on note maintenant que, grâce à la formule de dérivation des fonctions composées, on a que (cf. (2.10), § 2.2.2, en particulier) :

$$\varphi'(t) = (b - a) \cdot \nabla f\big(a + t(b - a)\big)$$

d'où

$$f(b) - f(a) = (b - a) \cdot \nabla f(c), \quad c = a + \theta(b - a).$$

En se rappelant que

$$|b - a| = \max_j |b_j - a_j|, \quad a = (a_1, \ldots, a_N), \; b = (b_1, \ldots, b_N),$$

on a

$$\big|f(b) - f(a)\big| = \left|\sum_{j=1}^{N} (b_j - a_j)\, \mathrm{D}_j f(c)\right| \leqslant \sum_{j=1}^{N} |b_j - a_j| \, \big|\mathrm{D}_j f(c)\big|$$

$$\leqslant |b - a| \sum_{j=1}^{N} \big|\mathrm{D}_j f(c)\big| \leqslant K|b - a|.$$

(ii) En appliquant l'inégalité de (i) sur chacune des fonctions f_i on aura

$$\big|f_i(b) - f_i(a)\big| \leqslant K|b - a|$$

d'où

$$\big|f(b) - f(a)\big| = \max_{i=1, \ldots, M} \big|f_i(b) - f_i(a)\big| \leqslant K|b - a| \qquad \blacklozenge$$

2.3.2 Corollaire

Si Ω est un domaine de $\mathbb{R}^N$ et $f : \Omega \to \mathbb{R}^M$ est une fonction dérivable telle que $f'(x) = 0$, $x \in \Omega$ alors f est constante.

DÉMONSTRATION

L'hypothèse $f'(x) = 0$, $x \in \Omega$ équivaut à la condition que $D_j f_i(x) = 0$ pour tout $x \in \Omega$, $i = 1, \ldots, M$, $j = 1, \ldots, N$, où $f = (f_1, \ldots, f_M)$. En appliquant (ii) (prop. 2.3.1) à f restreinte à une boule ouverte contenue dans Ω, on conclut que f est localement constante dans Ω ; comme Ω est un domaine (donc connexe) on en déduit que f est constante dans Ω (prop. 1.8.2(v)). ♦

2.3.3 Corollaire

Soit Ω un ensemble ouvert et convexe dans $\mathbb{R}^N$; si $f : \Omega \to \mathbb{R}^M$ est une fonction dérivable telle que

$$K = \sup_{x \in \Omega} \|f'(x)\| = \sup \left\{ \sum_{j=1}^{N} |D_j f_i(x)| : i = 1, \ldots, M,\ x \in \Omega \right\} < \infty$$

où $f = (f_1, \ldots, f_M)$ alors

$$|f(x) - f(y)| \leqslant K|x - y|, \quad x, y \text{ dans } \Omega.$$

(Autrement dit, alors f est *lipschitzienne* dans Ω, § 2.3.4.) C'est une conséquence immédiate de (ii) de la proposition 2.3.1.

2.3.4 Remarques

Il faut remarquer que l'analogue exact de (i) de la proposition 2.3.1 pour les fonctions vectorielles $f : \Omega \to \mathbb{R}^M$ est en défaut (Ex. 2, § 2.3.5). Si $f = (f_1, \ldots, f_M)$, il est bien vrai que pour chaque $i = 1, \ldots, M$, il existe un c_i dans le segment $]a, b[$ tel que

$$f_i(b) - f_i(a) = (b - a) \cdot \nabla f_i(c_i) ;$$

mais on ne peut pas toujours faire en sorte que les c_i soit égaux ; autrement dit, une formule du style

$$f(b) - f(a) = f'(c)(b - a)$$

pour un choix approprié de c dans $]a, b[$ n'est pas toujours possible pour le cas $f : \Omega \to \mathbb{R}^M$.

Néanmoins, on peut écrire une formule valable pour toutes les fonctions $f : \Omega \to \mathbb{R}^M$ qui *sont de classe C^1* (en supposant toujours que le segment $[a, b]$ soit contenu dans Ω). Supposons que $f = (f_1, \ldots, f_M)$; formons la fonction φ exactement comme dans la démonstration de la proposition 2.3.1 :

$$\varphi(t) = f\big(a + t(b - a)\big) = \big(\varphi_1(t), \ldots, \varphi_M(t)\big), \quad 0 \leqslant t \leqslant 1.$$

En utilisant la forme intégrale de la formule des accroissements finis (indiquée au début de cette section) on aura, pour $i = 1, \ldots, M$,

$$f_i(b) - f_i(a) = \varphi_i(1) - \varphi_i(0) = \int_0^1 \varphi_i'(t)\,\mathrm{d}t$$
$$= \int_0^1 \{\nabla f_i(a + t(b-a)) \cdot (b-a)\}\,\mathrm{d}t$$

ce que nous pouvons écrire comme

$$f_i(b) - f_i(a) = \left(\int_0^1 \nabla f_i(a + t(b-a))\,\mathrm{d}t\right) \cdot (b-a)$$

où

$$\int_0^1 \nabla f_i(a + t(b-a))\,\mathrm{d}t$$
$$= \left(\int_0^1 \mathrm{D}_1 f_i(a + t(b-a))\,\mathrm{d}t, \ldots, \int_0^1 \mathrm{D}_N f_i(a + t(b-a))\,\mathrm{d}t\right).$$

D'ici, nous voyons que nous pouvons écrire la formule

$$f(b) - f(a) = \left(\int_0^1 f'(a + t(b-a))\,\mathrm{d}t\right)(b-a) \tag{2.11}$$

où

$$\int_0^1 f'(a + t(b-a))\,\mathrm{d}t = \left[\int_0^1 \mathrm{D}_j f_i(a + t(b-a))\,\mathrm{d}t\right]_{\substack{i=1,\ldots,M \\ j=1,\ldots,N}}$$

est la matrice de type (M, N) dont le terme (i, j) est obtenu en intégrant $\mathrm{D}_j f_i$ sur le segment $[a, b]$ selon la formule explicitée. La formule (2.11) peut être considérée comme la forme intégrale de la formule des accroissements finis dans le cas général d'une fonction f de type $\mathbb{R}^N \to \mathbb{R}^M$.

Les fonctions satisfaisant une inégalité du type indiqué dans le corollaire 2.3.3 s'appellent les *fonctions lipschitziennes*. Plus précisément, soit $f : \Omega \to \mathbb{R}^M$ une fonction définie dans un sous-ensemble *quelconque* Ω de $\mathbb{R}^N$; on dit que f est ***lipschitzienne*** dans Ω s'il existe une constante $K > 0$ telle que

$$|f(x) - f(y)| \leqslant K|x - y|\,, \quad x, y \in \Omega\,.$$

On peut alors formuler le corollaire 2.3.3 comme suit : *une fonction dérivable $f : \Omega \to \mathbb{R}^M$ où Ω est un ensemble ouvert et connexe dans $\mathbb{R}^N$ est lipschitzienne si*

$$\sup_{x \in \Omega} \|f'(x)\| < \infty\,.$$

Pour une réciproque de cet énoncé voir l'exercice 5 du paragraphe 2.3.5.

Du corollaire 2.3.3, on déduit, avec un raisonnement supplémentaire, l'énoncé suivant : *si* $f \in C^1(\Omega;\mathbb{R}^M)$, Ω *étant un ensemble ouvert quelconque de* $\mathbb{R}^N$, *alors, pour toute partie compacte* A *de* Ω, *la restriction de* f *à* A *est une application lipschitzienne*. En effet, si $f = (f_1, \ldots, f_M)$, la continuité des fonctions $\mathrm{D}_j f_i$ dans Ω entraîne qu'elles sont bornées dans toute partie compacte de Ω ; selon le corollaire 2.3.3, f est donc lipschitzienne dans toute boule fermée contenue dans Ω. Si f n'était pas lipschitzienne dans A, on aurait des points ξ_n, η_n, $n = 1, 2, \ldots$, dans A tels que

$$\left|f(\xi_n) - f(\eta_n)\right| > n|\xi_n - \eta_n| \,; \tag{2.12}$$

comme A est compact, on peut supposer, en passant à une sous-suite appropriée, que ξ_n converge vers un point ξ de A lorsque $n \to \infty$; si $\left|f(x)\right|$ est majoré par C pour $x \in A$ on aura, de (2.12), que

$$|\xi_n - \eta_n| < \frac{2C}{n} \,,$$

d'où on déduit que η_n aussi converge vers ξ lorsque $n \to \infty$. Si $\overline{B}(\xi,\varepsilon) = P$ est contenue dans Ω, alors f sera lipschitzienne dans la boule fermée P ; or ceci contredit l'inégalité (2.12) car ξ_n, η_n seront dans P pour tout n suffisamment grand. Donc f est lipschitzienne dans A.

Il faut remarquer qu'une restriction du type « convexité » imposée sur Ω du corollaire 2.3.3 est nécessaire pour la validité de sa conclusion ; en effet, il existe des fonctions $f : \Omega \to \mathbb{R}$ de classe C^1 (Ω étant un ensemble ouvert de $\mathbb{R}^N$) telles que $|\mathrm{D}_j f|$, $j = 1, \ldots, N$, sont bornées dans Ω sans que f soit lipschitzienne dans Ω. Un exemple d'une telle fonction est donnée par

$$f(x_1,x_2) = \arg(x_1,x_2) \,, \quad (x_1,x_2) \in \Omega = \mathbb{R}^2 \setminus \{(x_1,0) : x_1 \leqslant 0\} \,;$$

cet exemple est développé en détail dans le chapitre 3 ; intuitivement, $\arg(x_1,x_2)$ est l'angle θ, pris entre $-\pi$ et π, dans l'écriture polaire

$$x_1 = r\cos\theta \,, \quad x_2 = r\sin\theta \,.$$

Une fonction $f : \Omega \to \mathbb{R}^M$, Ω étant une partie quelconque de $\mathbb{R}^N$, s'appelle ***localement lipschitzienne*** dans Ω si, pour tout point a de Ω, il existe une boule $B(a,\varepsilon)$ telle que la restriction de f à $B(a,\varepsilon) \cap \Omega$ est lipschitzienne ; autrement dit, il existe un $K > 0$ et un $\varepsilon > 0$ (dépendants de $a \in \Omega$) tels que

$$\left|f(x) - f(y)\right| \leqslant K|x - y| \,, \quad x, y \in B(a;\varepsilon) \cap \Omega \,.$$

On voit immédiatement du corollaire 2.3.3 que *toute fonction* f *de* $C^1(\Omega;\mathbb{R}^M)$, Ω *ouvert dans* $\mathbb{R}^N$, *est localement lipschitzienne*.

En utilisant les relations entre les normes $|\cdot|$ et $\|\cdot\|$, on voit immédiatement que dans la définition de la propriété lipschitzienne on peut remplacer la norme $|\cdot|$ par la norme euclidienne $\|\cdot\|$; de même, dans les inégalités de la proposition 2.3.1 ou du corollaire 2.3.3, on peut utiliser la norme $\|\cdot\|$, quitte à remplacer K par $K\sqrt{M}$.

2.3.5 Exercices

[1] Soit $f(x,y) = xy$, $(x,y) \in \mathbb{R}^2$; pour $a = (0,0)$, $b = (1,1)$, vérifier que le c de (i) de la proposition 2.3.1 est uniquement déterminé alors que pour $a = (1,0)$, $b = (1,1)$, le c en question n'est pas unique ; plus généralement, si $a = (x_0, y_0)$, $b = (x_1, y_1)$, alors le c est unique si et seulement si $(x_1 - x_0)(y_1 - y_0) \neq 0$.

[2] Soit $f(x) = (\cos x, \sin x)$, $x \in \mathbb{R}$; vérifier que $f(0) = f(2\pi)$ et $\|f'(x)\| = 1$; en conclure qu'il n'existe pas de c tel que $f(2\pi) - f(0) = f'(c)2\pi$. En se servant de (ii) de la proposition 2.3.1, vérifier que $|\sin x| \leqslant |x|$, $1 - \cos x \leqslant |x|$, $x \in \mathbb{R}$.

[3] Soit Ω un ensemble ouvert convexe dans $\mathbb{R}^N$. Si f, g sont deux fonctions dérivables définies dans Ω à valeurs dans $\mathbb{R}^M$ telles que $f'(x) = g'(x)$, $x \in \Omega$ alors $f - g$ est une constante. En prenant $g(x) = f(a) + f'(a)(x - a)$ où a est un point de Ω, conclure que si $f'(x)$ est constante dans Ω alors f est égale à la fonction affine g.

[4] Soit $\Omega \subset \mathbb{R}^N$. Une fonction $f : \Omega \to \mathbb{R}^M$ lipschitzienne est uniformément continue dans Ω.

[5] (Une réciproque du corollaire 2.3.3) Soit Ω une partie ouverte de $\mathbb{R}^N$; si $f : \Omega \to \mathbb{R}^M$ est une fonction dérivable lipschitzienne, démontrer que

$$\sup_{x \in \Omega} \|f'(x)\| < \infty .$$

(On a $|f(x+h) - f(x)| \leqslant K|h|$, $x \in \Omega$ et $f(x+h) = f(x) + f'(x)h + |h|\,\rho(x,h)$; en prenant $h = t\xi$, t réel, $t \to 0$, on obtient $|f'(x)\xi| \leqslant K$ d'où, avec $\xi = (1, 0, \ldots, 0)$ etc., on aura $|\mathrm{D}_j f_i(x)| \leqslant K$.)

[6] Soit $f : \Omega \to \mathbb{R}$, Ω ouvert dans $\mathbb{R}^N$, une fonction dérivable ; soit $\gamma : [0, 1] \to \Omega$ (un chemin) dérivable avec $\gamma(0) = a$, $\gamma(1) = b$. Démontrer qu'il existe un θ, $0 < \theta < 1$, tel que

$$f(b) - f(a) = f'\big(\gamma(\theta)\big)\gamma'(\theta) = \nabla f\big(\gamma(\theta)\big) \cdot \gamma'(\theta) .$$

Voir que, avec $\gamma(t) = a + t(b - a)$, on retrouve (i) de la proposition 2.3.1. Si $f(a) = f(b)$ on conclut qu'il existe un point $\gamma(\theta)$ du chemin γ où le gradient de f est orthogonal à la direction tangente au chemin à ce point.

7 Soit $f \in C^1(\Omega)$ où Ω est une partie ouverte et connexe de $\mathbb{R}^N$. Pour tout $a \in \Omega$ il existe une fonction $g \in C(\Omega; \mathbb{R}^N)$, $g = (g_1, \ldots, g_N)$, telle que

$$f(x) = f(a) + g(x) \cdot (x - a), \quad x \in \Omega.$$

En fait,

$$g_i(x) = \int_0^1 \mathrm{D}_i f\big(a + t(x - a)\big)\, \mathrm{d}t, \quad i = 1, \ldots, N.$$

(Utiliser (2.11) pour le cas $M = 1$; noter que, les $\mathrm{D}_i f$ étant continues, les g_i sont continues.)

2.4 MAXIMUM ET MINIMUM

Soit f une fonction réelle définie dans une partie Ω de $\mathbb{R}^N$. Un point $a \in \Omega$ s'appelle un point de ***maximum relatif*** pour f s'il existe un voisinage ouvert V de a tel que

$$f(x) \leqslant f(a), \quad x \in V \cap \Omega ;$$

s'il y a inégalité stricte ci-dessus, pour $x \neq a$, on dit que a est un point de ***maximum relatif strict***. On dit que a est un point de ***maximum absolu*** si

$$f(x) \leqslant f(a), \quad x \in \Omega.$$

Evidemment, un point de maximum absolu est un point de maximum relatif, mais la réciproque n'est pas nécessairement vraie. Les définitions concernant le ***minimum relatif***, le ***minimum relatif strict*** et le ***minimum absolu*** sont obtenues en renversant les inégalités ; par ailleurs, f admet un minimum relatif (respectivement, minimum relatif strict ou minimum absolu) au point a si et seulement si $-f$ admet un maximum relatif (respectivement, maximum relatif strict ou maximum absolu) au point a. Par ***extremum relatif*** (***strict***) on entend un maximum ou minimum relatif (strict). Les expressions ***maximum local***, ***minimum local***, etc., au lieu de maximum relatif, minimum relatif, etc., sont aussi d'usage courant.

La proposition simple suivante est utile pour la détermination des extremum relatifs.

2.4.1 Proposition (*condition nécessaire de premier ordre pour un extremum relatif*)

Soit $f : \Omega \to \mathbb{R}$ dérivable, Ω étant une partie ouverte de $\mathbb{R}^N$. Si $a \in \Omega$ est un point d'extremum relatif pour f alors

$$\mathrm{D}_j f(x) = 0, \quad j = 1, \ldots, N ;$$

autrement dit, $f'(a) = 0$ ou grad $f(a) = 0$.

DÉMONSTRATION

Considérons la fonction d'une variable réelle t définie par

$$\varphi(t) = f(a + te_j)$$

où $e_j = (0, \ldots, 1, \ldots, 0)$ est le j-ième élément de la base canonique de $\mathbb{R}^N$; Ω étant ouvert, φ est définie, au moins, dans un intervalle ouvert contenant 0. Si a est un point de maximum ou minimum relatif pour f, 0 est un point de maximum ou minimum relatif pour φ ; comme φ est dérivable, à cause de la dérivabilité de f, on aura $\varphi'(0) = 0$ ce qui veut dire que $\mathrm{D}_j f(a) = 0$. ♦

2.4.2 Remarques

Un point a où $f'(a) = 0$ s'appelle un point ***critique*** de f ; dans quelques autres contextes, un tel point s'appelle aussi un point de ***stationnarité*** de f ; on dit aussi que f est ***stationnaire*** au point a.

Nous savons à partir des exemples simples de fonctions d'une variable réelle qu'un point critique n'est pas nécessairement un point d'extremum relatif. Ainsi, des deux fonctions de $x \in \mathbb{R}$

$$g(x) = x^2\,, \quad h(x) = x^3$$

g possède un minimum relatif (même minimum absolu strict) au point critique $x = 0$ alors que pour h, $x = 0$ est un point critique sans être un minimum ou maximum relatif. Nous étudierons les points critiques en utilisants les dérivées de second ordre dans le chapitre 3.

La proposition simple 2.4.1 nous permet souvent de déterminer aussi les points de maximum ou minimum absolus. Supposons, par exemple, qu'il faut trouver un point de maximum absolu pour une fonction réelle, dérivable, non constante, f définie partout dans $\mathbb{R}^N$; si $f \geqslant 0$ et $f(x) \to 0$ lorsque $|x| \to \infty$ alors il est certain que le supremum de f dans $\mathbb{R}^N$ est atteint à un point a :

$$f(a) = \sup_{x \in \mathbb{R}^N} f(x)\,.$$

En effet, si $p \in \mathbb{R}^N$ est tel que $f(p) > 0$, on peut choisir un $R > 0$ tel que

$$f(x) < f(p)\,, \quad |x| > R\,;$$

on en conclut que

$$\sup_{x \in \mathbb{R}^N} f(x) = \sup_{|x| \leqslant R} f(x) = f(a)$$

pour un a tel que $|a| \leqslant R$, en se servant de la proposition qu'une fonction réelle, continue dans un ensemble compact ($\{x : |x| \leqslant R\}$) atteint son supremum. Pour trouver le point a, il suffit maintenant de déterminer

$$E = \{x : \mathrm{D}_j f(x) = 0\,,\ j = 1, \ldots, N\}$$

et obtenir un a tel que

$$f(a) = \sup_{x \in E} f(x) .$$

Si E est un ensemble fini cette dernière étape est, en principe, élémentaire ; si E se réduit à un point alors ce dernier est le point a cherché. Remarquons que, sous les conditions imposées sur f, E ne peut pas être vide.

La proposition 2.4.1 est déduite de l'observation simple suivante : si le point a est un point d'extremum relatif pour f, alors la restriction de f à chaque droite passant par a possède aussi un extremum relatif au point a. Nous donnons un exemple illustratif pour montrer que la restriction de f à chaque droite passant par a peut bien avoir un extremum (minimum) relatif au point a sans que a soit un extremum relatif pour f. Soit

$$f(x,y) = (x^2 - y)(2x^2 - y) , \quad (x,y) \in \mathbb{R}^2 .$$

On vérifie que les équations

$$D_1 f(x,y) = 0 , \quad D_2 f(x,y) = 0$$

correspondent ici à

$$2x(4x^2 - 3y) = 0 , \quad -3x^2 + 2y = 0$$

possédant l'unique solution $x = y = 0$. Donc $(0,0)$ est le seul point critique de f ; mais $(0,0)$ n'est pas un extremum relatif pour f car $f(0,0) = 0$ et, dans chaque disque ouvert centré à $(0,0)$,

$$f(x,y) < 0 \text{ si } x^2 < y < 2x^2 , \quad f(x,y) > 0 \text{ si } y < 0 ,$$

de sorte que $f(x,y)$ change de signe dans tout voisinage ouvert de $(0,0)$. Néanmoins, la restriction de f à la droite d'équation $x = 0$ donne

$$f(0,y) = y^2 > 0 \quad \text{si } y \neq 0$$

et la restriction de f aux droites d'équation de la forme $y = mx$ donne

$$f(x,mx) = x^2(x - m)(2x - m)$$

ce qui est strictement positif pour $x \neq 0$ dans un intervalle autour de 0 ; ainsi, $(0,0)$ est un minimum relatif pour chacune de ces fonctions restreintes.

2.4.3 Un problème d'extremum absolu dans $\mathbb{R}^N$

Soient $\xi_1, \ldots, \xi_n$, n points, non nécessairement distincts, dans $\mathbb{R}^N$; à déterminer un point $a \in \mathbb{R}^N$ tel que la somme des carrés des distances euclidiennes entre ξ_i et a, $i = 1, \ldots, n$, soit un minimum.

Démontrons que *l'unique solution est donnée par*

$$a = \frac{1}{n}\sum_{i=1}^{n} \xi_i \,,$$

la ***moyenne*** où le ***barycentre*** de $\xi_1, \ldots, \xi_n$.

Le calcul différentiel n'est aucunement nécessaire pour obtenir la solution ; en posant

$$f(x) = \sum_{i=1}^{n} \|\xi_i - x\|^2 \,, \quad x \in \mathbb{R}^N \,,$$

il suffit de vérifier l'identité algébrique suivante :

$$f(x) = f(a) + N\|a - x\|^2 \,, \quad x \in \mathbb{R}^N \,;$$

il résulte que a est l'unique point de $\mathbb{R}^N$ tel que $\inf_x f(x) = f(a)$. En effet,

$$\|\xi_i - x\|^2 = \|\xi_i - a + a - x\|^2 = \|\xi_i - a\|^2 + \|a - x\|^2 + 2(\xi_i - a)\cdot(a - x)$$

d'où la sommation sur $i = 1, \ldots, n$ donne l'identité indiquée car

$$\sum_{i=1}^{n} (\xi_i - a) = 0 \,.$$

Notons que, si

$$x = (x_1, \ldots, x_N) \,, \quad \xi_i = (\xi_{i1}, \ldots, \xi_{iN}) \,,$$

alors

$$f(x) = \sum_{i=1}^{n}\sum_{j=1}^{N} (\xi_{ij} - x_j)^2$$

et

$$\mathrm{D}_j f(x) = -2\sum_{i=1}^{n} (\xi_{ij} - x_j) = 0 \,, \quad j = 1, \ldots, N$$

donne

$$x_j = \frac{1}{n}\sum_{i=1}^{n} \xi_{ij} \,, \quad j = 1, \ldots, N \,.$$

Donc $f'(x) = 0$ si et seulement si $x = a$, la moyenne des ξ_i.

Dans ce problème, la fonction dérivable f a un unique point critique qui est un point de minimum absolu pour f.

Nous étudions maintenant un problème analogue, apparemment plus simple mais en fait d'une nature très différente.

2.4.4 Un problème d'extremum absolu dans $\mathbb{R}^2$

Soient P_1, P_2, P_3, trois points non colinéaires dans $\mathbb{R}^2$; à déterminer un point $M \in \mathbb{R}^2$ tel que la somme des distances euclidiennes entre P_i et M, $i = 1, 2, 3$, soit un minimum.

La *solution* de ce problème est la suivante. Le point M est uniquement déterminé ; si T est l'ensemble fermé triangulaire de sommets P_1, P_2, P_3, alors la position de M est fixée par T comme suit.

Cas (i) Les trois angles aux sommets de T sont tous inférieurs à $2\pi/3$; dans ce cas, M est l'unique point à l'intérieur de T tel que chacun des angles

$$\angle P_1MP_2\,, \quad \angle P_1MP_3\,, \quad \angle P_2MP_3$$

est égal à $2\pi/3$.

Cas (ii) Un des trois angles aux sommets de T est plus grand ou égal à $2\pi/3$; dans ce cas, M coïncide avec le sommet de T qui contient cet angle.

Une solution élémentaire, se basant des constructions de la géométrie euclidienne plane, est possible ; mais, ici, nous allons suivre une voie analytique dans le but d'illustrer la proposition 2.4.1 — tout en nous servant du langage de la géométrie euclidienne plane, supposé connu.

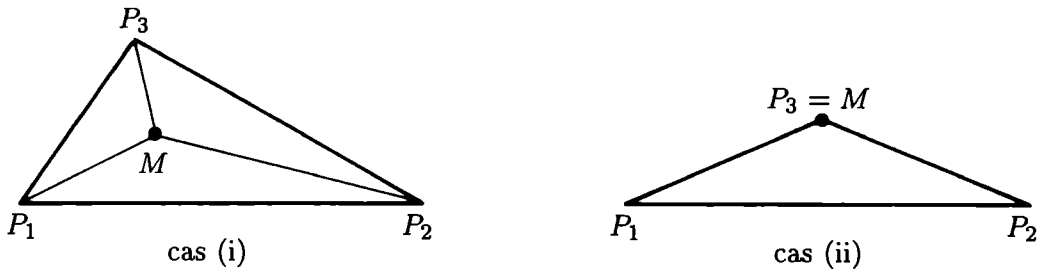


Fig. 2.1

Soient $P_i = (a_i, b_i)$, $i = 1, 2, 3$, $P = (x, y)$; alors, analytiquement, le problème revient à minimiser la fonction f définie par

$$f(P) = f(x,y) = \sum_{i=1}^{3} \|P - P_i\| = \sum_{i=1}^{3} \{(x-a_i)^2 + (y-b_i)^2\}^{1/2}\,;$$

plus précisément, il faut déterminer $M = (a, b)$ tel que

$$f(M) = \inf_{P \in \mathbb{R}^2} f(P) = \mu\,.$$

Rappelons que la description analytique de l'ensemble fermé triangulaire T est

$$T = \left\{ P = \sum_{i=1}^{3} \alpha_i P_i : \alpha_i \geqslant 0\,,\ i = 1, 2, 3\,,\ \sum_{i=1}^{3} \alpha_i = 1 \right\}\ ;$$

l'intérieur de T est formé des points P de T pour lesquels les $\alpha_i > 0$; le côté P_iP_j $(i \neq j)$ correspond aux points P de T tels que $\alpha_i + \alpha_j = 1$, et le sommet P_i s'obtient en prenant $\alpha_i = 1$, $\alpha_j = 0$, $j \neq i$.

Notons que f est une fonction continue partout ; elle est dérivable partout, *à l'exception des points* $P = P_i$, $i = 1, 2, 3$.

Montrons d'abord l'existence de M par un raisonnement général déjà esquissé dans le paragraphe 2.4.2. Comme

$$0 < f(P) \longrightarrow \infty \quad \text{lorsque } \|P\| \to \infty\,,$$

il est clair qu'il existe un disque fermé D, contenant T, tel que

$$\mu = \inf_{P \in D} f(P)\ ;$$

il suffit de choisir D comme un disque fermé, centré à l'origine, d'un rayon si grand que D contient T et que si $P \in D^c$, on a

$$f(P) > f(P_1)\,.$$

Comme D est un ensemble compact, l'infimum μ de f est atteint à au moins un point M de D ; d'ailleurs, tous les points minimisants M seront dans D.

Déterminons maintenant les points critiques de f ; nous allons voir que, dans le cas (i), f possède un unique point critique Q et que, dans le cas (ii), f n'a pas de point critique. D'ici, on obtiendra la solution complète déjà indiquée.

Un calcul montre que si $P = (x, y) \neq P_i$, $i = 1, 2, 3$, alors

$$\mathrm{D}_1 f(x,y) = \sum_{i=1}^{3} \frac{x - a_i}{\|P - P_i\|}\,, \quad \mathrm{D}_2 f(x,y) = \sum_{i=1}^{3} \frac{y - b_i}{\|P - P_i\|}\,.$$

En posant

$$\xi_i = \frac{P_i - P}{\|P_i - P\|} = \frac{1}{\|P_i - P\|}\,(a_i - x\,,\ b_i - x)\,,$$

il vient que $P = (x, y)$ est un point critique de f si et seulement si

$$\xi_1 + \xi_3 + \xi_3 = 0\,.$$

Notons que $\|\xi_i\| = 1$, $i = 1, 2, 3$; ainsi, si $i \neq j$,

$$\xi_i \cdot \xi_j = \cos\theta_{ij}\,, \quad 0 \leqslant \theta_{ij} \leqslant \pi\ (\theta_{ij} = \theta_{ji})\,,$$

où θ_{ij} correspond à l'angle $\angle P_iPP_j$. Vérifions que chaque θ_{ij} est égal à $2\pi/3$ (ce qui est évident à partir d'un théorème de statique bien connu concernant

l'équilibre des forces). Analytiquement, en prenant le produit scalaire de $\xi_1 + \xi_2 + \xi_3$ avec ξ_1, ξ_2, ξ_3 successivement, il vient que

$$-1 = \cos\theta_{12} + \cos\theta_{13} = \cos\theta_{12} + \cos\theta_{23} = \cos\theta_{13} + \cos\theta_{23}$$

d'où

$$\cos\theta_{12} = \cos\theta_{13} = \cos\theta_{23} = -\frac{1}{2}$$

ce qui donne $\theta_{ij} = 2\pi/3$.

Le calcul précédent montre que P est un point critique pour f si et seulement si P est un point tel que chaque angle $\angle P_iPP_j$ vaut $2\pi/3$. Un raisonnement géométrique établit maintenant que l'existence d'un tel point P est possible exactement dans le cas (i) ci-dessus — et alors un tel P est unique. En effet, on sait de la géométrie élémentaire que tous les points P, dans le demi-plan déterminé par le côté P_1P_3 et contenant P_2, tels que l'angle $\angle P_1PP_3$ est constant ($= 2\pi/3$ ici) forment un arc C d'un cercle ; de même, les points P, dans le demi-plan déterminé par le côté P_2P_3 et contenant P_1, tels que l'angle $\angle P_2PP_3$ est constamment égal à $2\pi/3$ forment un arc C' d'un autre cercle. L'intersection de ces deux arcs C, C' donnent, en plus du point P_3, un point P tel que chacun des angles

$$\angle P_1PP_3\,, \quad \angle P_2PP_3$$

vaut $2\pi/3$; si P est à l'intérieur de T alors l'angle $\angle P_1PP_3$ sera aussi égal à $2\pi/3$ et on aura le cas (i). Si P est à l'extérieur de T alors l'angle $\angle P_1PP_3$ sera inférieur à $2\pi/3$ et f n'a pas de point critique ; c'est le cas (ii).

Considérons maintenant le cas (i) ; on a vu qu'ici f possède un unique point critique Q, décrit ci-dessus, qui est à l'intérieur de T. Démontrons que Q est l'unique point M qui minimise f en établissant que

$$f(Q) < f(P_i)\,, \quad i = 1, 2, 3\,,$$

car alors, M, étant différent des P_i, sera un point critique de f ; l'unique point critique pour f étant Q, on aura $M = Q$. En reprenant la notation

$$\xi_i = \frac{P_i - Q}{\|P_i - Q\|}\,, \quad i = 1, 2, 3\,,$$

on aura

$$\begin{aligned}\|P_1 - P_3\| > (P_1 - P_3)\cdot\xi_1 &= (P_1 - Q)\cdot\xi_1 + (Q - P_3)\cdot\xi_1 \\ &= \|P_1 - Q\| - \|P_3 - Q\|\xi_3\cdot\xi_1\ ;\end{aligned}$$

d'où, en utilisant $\xi_1\cdot\xi_3 = -1/2$ (Q étant le point critique de f décrit ci-dessus), on aura

$$\|P_1 - P_3\| > \|P_1 - Q\| + \frac{1}{2}\|P_3 - Q\|\,.$$

De même,

$$\|P_2 - P_3\| > \|P_2 - Q\| + \frac{1}{2}\|P_3 - Q\|$$

d'où

$$f(P_3) > f(Q)\,.$$

On établit les deux inégalités

$$f(P_1) > f(Q)\,, \quad f(P_2) > f(Q)$$

de la même manière, complétant ainsi la solution dans le cas (i).

Considérons maintenant le cas (ii) ; supposons que l'angle qui est supérieur ou égal à $2\pi/3$ se trouve au sommet P_3 ; montrons que dans ce cas $M = P_3$. En effet, dans ce cas, f ne possède pas de point critique et un point minimisant M doit être égal à l'un des points P_1, P_2, P_3. Comme l'angle obtus du triangle T se situe au sommet P_3, on aura

$$\|P_1 - P_2\| > \max\{\|P_1 - P_3\|\,, \|P_2 - P_3\|\}$$

d'où

$$f(P_3) = \|P_3 - P_1\| + \|P_3 - P_2\| < \|P_3 - P_1\| + \|P_2 - P_1\| = f(P_1)\,;$$

on a, de même,

$$f(P_3) < f(P_2)$$

de sorte que

$$f(P_3) = \min\{f(P_i) : i = 1, 2, 3\}\,,$$

ce qui donne P_3 comme l'unique point minimisant M. Ceci complète la solution du problème.

Le problème traité dans ce paragraphe était proposé par le mathématicien français Fermat (1601-1665) au physicien italien Torricelli (1608-1647), un élève célèbre de Galilée (1564-1642), qui l'a résolu de plusieurs manières. Une solution simple utilisant la géométrie élémentaire était donnée par Viviani (1622-1703), un mathématicien-physicien italien, élève de Galilée et Torricelli. On pourra trouver cette solution dans Dörrie [15 , pp. 361-363]. Différentes généralisations de ce problème étaient étudiées par le géomètre suisse Steiner (1796-1863) (cf. ses *Gesammelle Werke II*, *Chelsea Publishing Co.* pp. 16-17, pp. 93-95). Le problème de déterminer un point $M \in \mathbb{R}^N$ tel que la somme des distances euclidiennes entre P_i et M, $i = 1, \ldots, n$, soit un minimum ne semble pas posséder une solution simple. Déjà le cas $n = 4$ dans $\mathbb{R}^2$ présente une multitude de possibilités selon la disposition des quatre points dans $\mathbb{R}^2$.

2.4.5 Exercices

[1] Déterminer les points critiques pour chacune des fonctions suivantes et discuter leur extrémalité, $(x,y) \in \mathbb{R}^2$:

(i) $f(x,y) = x + y$;
(ii) $f(x,y) = x^2 + y^2$;
(iii) $f(x,y) = xy$;
(iv) $f(x,y) = x + y + xy$;
(v) $f(x,y) = (x^2 + y^2)\exp\{-(x^2 + y^2)\}$;
(vi) $f(x,y) = x^2 + 2y^2 + 2xy + 4y + 2$.

[2] Soit

$$f(x,y) = xy + \frac{a}{x} + \frac{b}{y}, \quad x \neq 0,\ y \neq 0,\ (x,y) \in \mathbb{R}^2,$$

où $a > 0$, $b > 0$; vérifier que $(a^{2/3}b^{-1/3}, a^{-1/3}b^{2/3})$ est le seul point critique pour f ; démontrer que ce point critique est un point de minimum relatif en raisonnant que, pour tout $\varepsilon > 0$, $\inf\{f(x,y) : x > \varepsilon,\ y > \varepsilon\}$ est atteint.

[3] Soit $f(x,y) = (x^2 - y)(ax^2 - y)$, $(x,y) \in \mathbb{R}^2$ $(a \in \mathbb{R})$. Si $a \neq 1$, le seul point critique de f est $(0,0)$ qui n'est pas un point d'extremum relatif. Si $a = 1$, les points critiques sont $\{(x,y) : y = x^2\}$; ils sont tous les points de minimum absolu.

[4] Soit $f(x,y) = ax^2 + 2bxy + cy^2$, $(x,y) \in \mathbb{R}^2$, $(a,b,c) \neq (0,0,0)$; posons $D = ac - b^2$. Vérifier que si $D \neq 0$, alors $(0,0)$ est le seul point critique pour f ; si $D > 0$, $a > 0$, on a $f(x,y) > 0$ pour $(x,y) \neq (0,0)$ de sorte que $(0,0)$ est alors un point de minimum absolu pour f (écrire f dans la forme $\left(\sqrt{a}x + (b/\sqrt{a})\,y\right)^2 + (ac - b^2/a)\,y^2\big)$; si $D > 0$, $a < 0$, alors $f(x,y) < 0$ pour $(x,y) \neq (0,0)$ de sorte que $(0,0)$ est dans ce cas un point de maximum absolu pour f ; si $D < 0$ alors $(0,0)$ n'est pas un extremum relatif. Si $D = 0$ alors les points critiques de f forment une droite passant par $(0,0)$ (d'équation $ax + by = 0$ si $b \neq 0$) ; dans ce cas $f(x,y)$ est de la forme $\pm(\alpha x + \beta y)^2$; les points critiques sont des points de minimum absolu si le signe est $+$, de maximum absolu si le signe est $-$.

2.5 CONVERGENCE DES DÉRIVÉES

Une question naturelle est la suivante : si une suite de fonctions dérivables f_n converge vers une fonction f lorsque $n \to \infty$, est-il vrai que f est dérivable et, si c'est le cas, est-ce que f_n' converge vers f' lorsque $n \to \infty$? Commençons par deux exemples simples qui montrent que la question, dans cette forme, ne

peut pas avoir une réponse affirmative, même si la convergence de f_n vers f est uniforme.

Soit

$$f_n(x) = \frac{\sin nx}{n}, \quad n = 1, 2, \ldots, \ x \in \mathbb{R} ;$$

alors, pour tout $x \in \mathbb{R}$, $f_n(x) \to f(x) = 0$ lorsque $n \to \infty$, uniformément en $x \in \mathbb{R}$ car

$$|f_n(x)| \leqslant \frac{1}{n}, \quad x \in \mathbb{R}.$$

Mais $f_n'(x) = \cos nx$ converge vers 1 si x est un multiple de 2π (en fait, $f_n'(2k\pi) = 1$, $k \in \mathbb{Z}$), alors que $f'(x) \equiv 0$; si $x = \pi/2$, la suite $f_n'(x) = \cos n\pi/2$ prend chacune des valeurs 0, +1, −1 infiniment souvent de sorte que $f_n'(\pi/2)$ ne converge vers rien ; en fait, la suite $f_n'(x) = \cos nx$ est une suite non convergente dès que x est un nombre qui n'est pas un multiple entier de 2π ; mais nous n'avons pas besoin de ce fait (dont la démonstration n'est pas immédiate) pour ce que nous voulons illustrer ici.

Soit, pour $n = 2, 3, \ldots$,

$$f_n(x) = \begin{cases} nx^2 & \text{si } |x| \leqslant \dfrac{1}{n} \\ 2x - \dfrac{1}{n} & \text{si } \dfrac{1}{n} < x < 1 \\ -2x - \dfrac{1}{n} & \text{si } -1 < x < -\dfrac{1}{n}. \end{cases}$$

On vérifie que les f_n sont continûment dérivables dans $]-1, 1[$ et que

$$\lim_{n\to\infty} f_n(x) = f(x) = 2|x|, \quad -1 < x < 1$$

uniformément en x car

$$\sup_{-1<x<1} |f_n(x) - 2|x|| \leqslant \frac{3}{n} ;$$

ce dernier est une conséquence des estimations suivantes :

$$|f_n(x) - 2|x|| \leqslant \frac{1}{n} \quad \text{si } |x| > \frac{1}{n}$$

$$|f_n(x) - 2|x|| \leqslant |f_n(x)| + 2|x| \leqslant \frac{1}{n} + \frac{2}{n} \quad \text{si } |x| \leqslant \frac{1}{n}.$$

Dans cet exemple, la suite de fonctions dérivables f_n converge uniformément (lorsque $n \to \infty$) vers une fonction f qui n'est pas dérivable partout.

La proposition qui suit donne une réponse utile à une variation de la question posée ci-dessus.

2.5.1 Proposition

Soit $f_n : \Omega \to \mathbb{R}^M$, $n = 1, 2, \ldots$, *une suite de fonctions dérivables, définies dans un domaine* Ω *de* $\mathbb{R}^N$. *Supposons que :*

(i) il existe un point $\xi \in \Omega$ *tel que la suite* $\{f_n(\xi)\}$ *converge ;*
(ii) pour tout point $a \in \Omega$ *il existe un nombre* $\rho(a) > 0$ *tel que*

$$V_a = \{x : |x - a| < \rho(a)\} \subset \Omega$$

et tel que $\{\mathrm{D}_j f_{ni}\}_{n\in\mathbb{N}^*}$ *converge uniformément dans* V_a *où*

$$f_n = (f_{n1}, \ldots, f_{nM}), \quad i = 1, \ldots, M, \ j = 1, \ldots, N.$$

Alors,

$$\lim_{n\to\infty} f_n(x) = f(x), \quad x \in \Omega$$

existe, la convergence étant uniforme pour $x \in V_a$, *pour chaque* $a \in \Omega$ *; en posant*

$$f(x) = \big(f_1(x), \ldots, f_M(x)\big), \quad g_{ij}(x) = \lim_{n\to\infty} \mathrm{D}_j f_{ni}(x), \quad x \in \Omega,$$

on a que f *est dérivable dans* Ω *et* $\mathrm{D}_j f_i = g_{ij}$, $i = 1, \ldots, M$, $j = 1, \ldots, N$.

DÉMONSTRATION

Il suffit de considérer le cas $M = 1$, car le résultat général se déduit de ce cas en l'appliquant à chacune des M suites de fonctions réelles $\{f_{ni}\}$, $i = 1, \ldots, M$.

De la formule des accroissement finis (corollaire 2.3.3) on a, pour tout x, y dans V_a, $a \in \Omega$,

$$\big|f_n(x) - f_m(x) - \{f_n(y) - f_m(y)\}\big| \leqslant K_{mn}(a)|x - y| \leqslant 2K_{mn}(a)\,\rho(a) \quad (2.13)$$

où

$$K_{mn}(a) = \sup\left\{\sum_{j=1}^{N} \big|\mathrm{D}_j(f_n - f_m)(z)\big| : z \in V_a\right\}.$$

Selon l'hypothèse (ii) de la proposition, $K_{mn}(a) \to 0$ lorsque $m, n \to \infty$; on en conclut, à l'aide de (2.13), que si la suite $\{f_n(y)\}$ converge pour *un point* y de V_a, alors $\{f_n(x)\}$ converge pour *tout point* x de V_a et ceci uniformément en x dans V_a. Ceci établit aussi que l'ensemble

$$E = \Big\{x \in \Omega : \lim_{n\to\infty} f_n(x) \text{ existe}\Big\}$$

est à la fois ouvert et fermé dans Ω ; comme E est non vide selon l'hypothèse (i) et Ω est connexe on a que $\Omega = E$. Ainsi, la fonction f est définie dans Ω par

$$f(x) = \lim_{n\to\infty} f_n(x), \quad x \in \Omega.$$

Pour tout $a \in \Omega$, la convergence de f_n vers f étant uniforme dans V_a (voir ci-dessus), f est continue dans Ω. Montrons que f est dérivable à chaque point a de Ω et que

$$f'(a) = g(a) = \big(g_1(a), \ldots, g_N(a)\big)$$

où

$$g_j(a) = \lim_{n\to\infty} \mathrm{D}_j f_n(a)\,, \quad j = 1, \ldots, N\,.$$

Pour $a \in \Omega$ et $\varepsilon > 0$, choisissons un entier p tel que

$$K_{mn}(a) < \varepsilon\,, \quad \sum_{j=1}^{N} \big|\mathrm{D}_j f_n(a) - g_j(a)\big| < \varepsilon\,, \quad m \geqslant p\,,\ n \geqslant p\,, \tag{2.14}$$

ce qui est possible grâce à l'hypothèse (ii). Prenons $y = a$ dans (2.13) et laissons $m \to \infty$ dans (2.13) en fixant $n = p$; on obtient, grâce à (2.14), que

$$\big|f(x) - f(a) - \big\{f_p(x) - f_p(a)\big\}\big| \leqslant \varepsilon|x - a|\,, \quad x \in V_a\,. \tag{2.15}$$

La dérivabilité de f_p au point a donne l'existence d'un nombre r, $0 < r \leqslant \rho(a)$, tel que

$$\big|f_p(x) - f_p(a) - f_p'(a)\cdot(x - a)\big| \leqslant \varepsilon|x - a| \quad \text{si } |x - a| < r\,. \tag{2.16}$$

De (2.15) et (2.16) on obtient, pour $|x - a| < r$, que

$$\big|f(x) - f(a) - g(a)\cdot(x - a)\big| \leqslant 2\varepsilon|x - a| + \big|\big\{f_p'(a) - g(a)\big\}\cdot(x - a)\big| < 3\varepsilon|x - a|$$

en utilisant la deuxième inégalité dans (2.14) ; ceci établit que $f'(a)$ existe et est égal à $g(a)$. ♦

2.5.2 Corollaire

Soit $h_n : \Omega \to \mathbb{R}^M$, $n = 1, 2, \ldots$, une suite de fonctions dérivables, définies dans un domaine Ω de $\mathbb{R}^N$; posons

$$f_n = h_1 + \cdots + h_n\,.$$

Sous les hypothèses de la proposition 2.5.1 sur $\{f_n\}$, on a la convergence (uniforme pour $x \in V_a$) de la série $\sum_n h_n(x)$; en posant

$$f(x) = \sum_n h_n(x)\,, \quad h_n = \big(h_{n1}, \ldots, h_{nM}\big)\,, \quad f = (f_1, \ldots, f_M)\,;$$

on a, de plus, que f est dérivable dans Ω et que

$$\mathrm{D}_j f_i(x) = \sum_n \mathrm{D}_j h_{ni}(x)\,, \quad x \in \Omega\,,\ i = 1, \ldots, M\,,\ j = 1, \ldots, N\,.$$

La démonstration du corollaire est immédiate.

2.5.3 Remarques

Il est clair que si les fonctions f_n de la proposition 2.5.1 sont de classe C^1, alors la fonction limite f sera aussi de classe C^1 car les $\mathrm{D}_j f_i$, étant limites uniformes des $\mathrm{D}_j f_{ni}$ dans V_a, seront continues dans V_a pour tout a.

La proposition 2.5.1 s'énonce de manière plus courte et simple, si nous introduisons la notion de ***convergence uniforme locale*** de f_n' vers g dans Ω ; cette dernière veut dire simplement que pour tout $a \in \Omega$ il existe un voisinage ouvert U_a de a contenu dans Ω tel que

$$\sup_{x \in U_a} \left\| f_n'(x) - g(x) \right\| \longrightarrow 0 \quad \text{lorsque} \quad n \to \infty$$

où, pour une matrice $A = [a_{ij}]$, $\|A\| = \max_i \sum_j |a_{ij}|$, $g = [g_{ij}]$ et

$$\left\| f_n'(x) - g(x) \right\| = \max_i \sum_j \left| \mathrm{D}_j f_{ni}(x) - g_{ij}(x) \right| .$$

Comme tout voisinage ouvert U_a de a contenu dans Ω contient un voisinage ouvert du type V_a, il est évident que la proposition 2.5.1 peut être formulée comme suit :

Soit $f_n : \Omega \to \mathbb{R}^M$, $n = 1, 2, \ldots$, une suite de fonctions dérivables définies dans un domaine Ω de $\mathbb{R}^N$. Supposons que

(i) $\{f_n(\xi)\}$ converge pour un $\xi \in \Omega$,
(ii) il y a convergence uniforme locale de f_n' vers g dans Ω.

Alors f_n converge, lorsque $n \to \infty$, vers une fonction $f : \Omega \to \mathbb{R}^M$ qui est dérivable et $f' = g$; la convergence de f_n vers f est localement uniforme.

Le corollaire 2.5.2 s'énonce comme la validité des formules

$$f(x) = \sum_{n=1}^{\infty} h_n(x) , \quad f'(x) = \sum_{n=1}^{\infty} h_n'(x) , \quad x \in \Omega ,$$

sous les hypothèses appropriées sur $\sum_n h_n'(z)$ et la convergence de $\sum_n h_n(\xi)$ pour un ξ.

Rappelons ici un exemple, simple mais important, dû en principe à Weierstrass, qui montre de manière dramatique qu'une série convergente de fonctions dérivables ne peut pas être dérivée terme à terme. Soit f la fonction définie par

$$f(x) = \sum_{n=1}^{\infty} b^n \cos(a^n x) , \quad x \in \mathbb{R}, \tag{2.17}$$

où $0 < b < 1$, $a > 1$, tel que $ab \geqslant 1$. Alors $f \in C(\mathbb{R})$ mais f est non dérivable partout. La continuité de f est immédiate car l'inégalité

$$\left| b^n \cos a^n x \right| \leqslant b^n$$

entraîne que la série définissant f est normalement convergente dans $\mathbb{R}$ (§ 1.9.2). La démonstration exacte pour la non dérivabilité de la fonction f de (2.17) est longue ; Weierstrass et les auteurs de la plupart des livres courants considèrent seulement le cas où a est un entier impair suffisamment grand ; le résultat précis cité ci-dessus est dû à Hardy, qui a analysé cette fonction f en détail sous les hypothèses générales mentionnées précédemment (cf. son article de 1911 « *Weierstrass's non-differentiable function* » reproduit dans ses *Collected Papers*, vol. IV, pp. 477-501). L'existence des fonctions continues et nulle part dérivables fit une grande sensation au XIX^{e} siècle ; aujourd'hui, on sait que de telles fonctions sont très répandues (dans un sens précis) et on a plusieurs types d'exemples de fonctions continues non dérivables qui sont plus faciles à analyser que la fonction f de (2.17).

2.6 INTERPRÉTATIONS DE LA DÉRIVÉE

Dans cette section nous présentons quelques considérations géométriques et physiques en relation avec la notion de dérivée. La présentation sera quelquefois un peu informelle ; elle utilisera les notions intuitives de la géométrie comme celles de courbe, de surface, de tangente, etc. Certaines de ces notions seront reprises plus précisément dans des chapitres suivants.

Une fonction f de type $\mathbb{R}^N \to \mathbb{R}^M$ peut être interprétée de plusieurs façons différentes. Si $N = 1$, on peut penser à f comme décrivant la *trajectoire* d'un point matériel ou une *courbe* dans $\mathbb{R}^M$ (où M est souvent égal à 2 ou 3 dans les applications élémentaires). Les fonctions $f = (f_1, \ldots, f_M)$ de type $\mathbb{R} \to \mathbb{R}^M$ sont omniprésentes en toute théorie mathématique où on veut décrire la variation de M quantités $f_1, \ldots, f_M$ en fonction d'une variable réelle x, dite indépendante, interprétée selon la situation comme temps, volume, densité, température, etc. La dérivée

$$f'(x) = \big(f_1'(x), \ldots, f_M'(x)\big)$$

représente alors le *taux de variation infinitésimale* autour de x, exprimé souvent comme suit :

$$f_i(x + \mathrm{d}x) - f_i(x) = f_i'(x)\,\mathrm{d}x\,, \quad i = 1, \ldots, M\,.$$

Le vecteur $f'(x) = \big(f_1'(x), \ldots, f_M'(x)\big)$ est considéré comme le *vecteur de vitesse* au temps x si f est interprétée comme la trajectoire d'un point mobile ; il est interprété comme donnant la direction de la *tangente* au point $f(x)$ de la courbe décrite par f ; pour cette dernière interprétation, il est naturel de supposer que $f(x) \neq 0$ car le vecteur nul ne possède pas de direction.

Si f est une application de type $\mathbb{R}^2 \to \mathbb{R}$, on peut associer à f une *surface* dans $\mathbb{R}^3$ en considérant son graphe ; on dit que la surface est donnée par une équation cartésienne

$$z = f(x, y)\,, \quad (x, y) \in \Omega \subset \mathbb{R}^2\,,\ (x, y, z) \in \mathbb{R}^3\,.$$

Le gradient de f à un point (a,b) donne le plan tangent au point de coordonnées (a,b,c), $c = f(a,b)$, de la surface associé à f ; son équation sera

$$z - c = \mathrm{D}_1 f(a,b)(x-a) + \mathrm{D}_2 f(a,b)(y-b)\,.$$

Si le gradient de f au point (a,b) est nul on voit que le plan tangent est parallèle au plan $z = 0$ dans $\mathbb{R}^3$: son équation est

$$z - c = 0\,.$$

Ainsi, si (a,b) est un poïnt d'extremum relatif pour f, le plan tangent au point (a,b,c), $c = f(a,b)$, à la surface associée à f, sera parallèle au plan $z = 0$ dans $\mathbb{R}^3$; on décrit ceci en disant que le plan tangent est *horizontal*.

Associés à une fonction f de type $\mathbb{R}^2 \to \mathbb{R}$, on a des *ensembles de niveau*

$$E_c = \big\{(x,y) : f(x,y) = c\big\}\,.$$

On verra dans la suite que les E_c sont des courbes dans $\mathbb{R}^2$ si $\nabla f(x,y) \neq 0$ pour $(x,y) \in E_c$; en fait, $\nabla f(x,y)$ est la direction normale au point (x,y) de E_c. Les *courbes* (*lignes*) *de niveau* E_c permettent aussi de visualiser le comportement de f ; c'est un procédé familier dans les cartes géographiques ou météorologiques où ces lignes représentent des tracés pour lesquels une certaine caractéristique comme la pression, la température ou l'altitude est constante.

Une application f de type $\mathbb{R}^2 \to \mathbb{R}^3$ peut souvent être vue comme la donnée d'une surface *paramétrique* ayant les paramètres de description dans le domaine de définition de f ; pour cette interprétation, on constatera que la condition

$$\operatorname{rang} f'(x,y) = 2$$

est essentielle. On verra dans la suite que l'équation du plan tangent à un point $f(a,b)$ de la surface associée sera donnée par

$$\big\{\xi - f(a,b)\big\} \cdot \nu = 0\,, \quad \xi \in \mathbb{R}^3\,,$$

où

$$\nu = D_1 f(a,b) \times D_2 f(a,b)\,.$$

Lorsque $f(x,y)$ est de la forme

$$f(x,y) = \big(x\,,\, y\,,\, \varphi(x,y)\big)$$

où φ est une fonction de type $\mathbb{R}^2 \to \mathbb{R}$, on retrouve les résultats de la discussion donnée ci-dessus sur la surface associée au graphe de la fonction φ.

Une autre façon de décrire des surfaces dans $\mathbb{R}^3$ est de considérer les *ensembles de niveau* d'une fonction f de type $\mathbb{R}^3 \to \mathbb{R}$:

$$E_c = \big\{(x,y,z) : f(x,y,z) = c\big\}\,.$$

On verra que E_c sera une surface si $\nabla f(x,y,z) \neq 0$ pour tout $(x,y,z) \in E_c$; la direction normale à un point (x,y,z) de E_c sera donnée par $\nabla f(x,y,z)$.

Ainsi, on voit qu'une surface dans $\mathbb{R}^3$ peut être donnée par au moins un des trois procédés suivants :

- comme graphe d'une fonction de type $\mathbb{R}^2 \to \mathbb{R}$;
- comme une surface paramétrique donnée par une fonction de type $\mathbb{R}^2 \to \mathbb{R}^3$;
- comme un ensemble de niveau d'une fonction de type $\mathbb{R}^3 \to \mathbb{R}$.

Le plan tangent et sa direction normale à un point de surface sont chaque fois déterminés par une expression calculée à partir de la dérivée de la fonction associée à chaque procédé.

Plus généralement, une fonction f de type $\mathbb{R}^N \to \mathbb{R}^M$ décrit la variation de M quantités

$$(f_1, \ldots, f_M) = f$$

en fonction de N variables réelles $x = (x_1, \ldots, x_N)$; on dit aussi que f est une fonction vectorielle (à valeur dans $\mathbb{R}^M$) des N variables (ou paramètres) $x_1, \ldots, x_N$. La terminologie est très variée, dépendant beaucoup du domaine d'application. Ainsi, une fonction $f : \mathbb{R} \to \mathbb{R}^M$ peut être appelée un *signal* (ou un *bruit*) s'il s'agit de la transmission de l'information. En physique, une fonction de type $\mathbb{R}^N \to \mathbb{R}$ s'appelle souvent un *champ scalaire* et une fonction de type $\mathbb{R}^N \to \mathbb{R}^N$ un *champ vectoriel* (souvent avec $N = 2$ ou 3). Dans tous les cas, la dérivée

$$f'(x) = [\mathrm{D}_j f_i(x)]$$

permet d'étudier la variation infinitésimale locale de $f = (f_1, \ldots, f_M)$ comme fonction de $x \in \mathbb{R}^N$ dans le sens que

$$f(x + \mathrm{d}x) = f(x) + f'(x)\,\mathrm{d}x\,.$$

Si f est un champ vectoriel de type $\mathbb{R}^N \to \mathbb{R}^N$, on verra que la ***divergence*** de f joue un rôle important dans son étude : si $f = (f_1, \ldots, f_N)$, $x \in \mathbb{R}^N$,

$$\operatorname{div} f(x) = \operatorname{Tr} f'(x) = \sum_{i=1}^{N} \mathrm{D}_i f_i(x)\,. \tag{2.18}$$

Dans le cas particulier de $N = 3$, on associe à un champ vectoriel f de type $\mathbb{R}^3 \to \mathbb{R}^3$ un autre champ vectoriel qui s'appelle son ***champ rotationnel*** (ou simplement ***rotationnel*** de f) ; si $f = (f_1, f_2, f_3)$, $x \in (x_1, x_2, x_3)$, le rotationnel de f au point x, $\operatorname{rot} f(x)$, est calculé par le calcul formel du déterminant qui suit :

$$\operatorname{rot} f(x) = \begin{vmatrix} e_1 & e_2 & e_3 \\ D_1 & D_2 & D_3 \\ f_1(x) & f_2(x) & f_3(x) \end{vmatrix} \tag{2.19}$$

$$= \big(\mathrm{D}_2 f_3(x) - \mathrm{D}_3 f_2(x)\,,\, \mathrm{D}_3 f_1(x) - \mathrm{D}_1 f_3(x)\,,\, \mathrm{D}_1 f_2(x) - \mathrm{D}_2 f_1(x)\big)\,.$$

On a déjà vu le champ vectoriel ∇f associé à un champ scalaire f de type $\mathbb{R}^N \to \mathbb{R}$; en fait, $\nabla f(x)$ est une autre notation pour $f'(x)$; elle est plus populaire dans ce contexte ; elle se note comme grad f aussi. Les trois ***opérateurs différentiels*** grad, div, rot joueront un rôle central dans plusieurs chapitres qui suivent.

2.7 EXERCICES

1 Démontrer la dérivabilité de la fonction f donnée par $f(x,y) = xy$, $(x,y) \in \mathbb{R}^2$, en partant de la définition de dérivabilité.

2 Soit $f(x,y) = \{25 - x^2 - y^2\}^{1/2}$, $x^2 + y^2 < 25$. Calculer $D_\nu f(3,2)$ pour :

(i) ν vecteur unitaire correspondant à l'angle 60° ;
(ii) ν vecteur unitaire correspondant à $(-2,3)$.

Que peut-on dire de la croissance de f dans les directions spécifiées dans (i) et (ii) ?

3 Si $z = yf(x^2 - y^2)$ vérifier que

$$y\frac{\partial z}{\partial x} + x\frac{\partial z}{\partial y} = \frac{xz}{y}$$

en spécifiant les hypothèses sur f, x, y.

4 Si $v(x,y,z) = u(r)$, $r = (x^2 + y^2 + z^2)^{1/2}$, vérifier que

$$\left(\frac{\mathrm{d}u}{\mathrm{d}r}\right)^2 = \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial z}\right)^2$$

en spécifiant l'hypothèse à faire sur la fonction u et en donnant les arguments des deux côtés de cette équation.

5 Calculer la dérivée directionnelle de la fonction donnée par $f(x,y) = \ln(x^2 + y^2)^{1/2}$, $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, au point $(1,1)$ dans la direction de la bissectrice du premier quadrant.

6 Soit $f(x,y,z) = \ln(e^x + e^y + e^z)$; calculer $D_\nu f(0,0,0)$ où ν est le vecteur unitaire formant les angles α, β, γ avec les axes des coordonnées.

7 Soit $u(x,y) = y^2/x$, $x \neq 0$; vérifier que la dérivée de u en tout point de l'ellipse d'équation $2x^2 + y^2 = c^2$ dans la direction de la normale à l'ellipse est égale à zéro.

8 Soit $f(x) = (\|x\|^2 - 1)x$, $x \in \mathbb{R}^2$. Démontrer que $f'(x) \neq 0$ (matrice nulle de type (2,2)) pour $x \in \mathbb{R}^2$ alors que $f(x) = 0$ sur la sphère $S = \{x : \|x\| = 1\}$.

(Ceci donne un contre-exemple au «théorème de Rolle» dans le cas d'une fonction f de type $\mathbb{R}^2 \to \mathbb{R}^2$.)

9 Soit Ω un ensemble ouvert borné dans $\mathbb{R}^N$; soit $f : \overline{\Omega} \to \mathbb{R}$ continue, dérivable dans Ω, constante sur $\operatorname{Fr}\Omega$. Démontrer qu'il existe un point $c \in \Omega$ tel que $\nabla f(c) = 0$ (théorème de Rolle). (Etudier le point c où f atteint un extremum global.)

10 Soit Ω un ensemble ouvert borné dans $\mathbb{R}^N$; soit $f : \overline{\Omega} \to \mathbb{R}^M$ continue, dérivable dans Ω. Supposons qu'il existe un $\xi \in \mathbb{R}^M$ $(\xi \neq 0)$ tel que $f(x) \cdot \xi$ est constante sur $\operatorname{Fr}\Omega$. Démontrer qu'il existe un $c \in \Omega$ tel que $f'(c)h \cdot \xi = 0$, $h \in \mathbb{R}^N$, c'est-à-dire ${}^t\xi f'(c) = 0$ en considérant $f'(c)$ comme une matrice de type (M, N) et ${}^t\xi$ comme une matrice ligne de type $(1, M)$. (Considérer les extremums de la fonction réelle g donnée par $g(x) = f(x) \cdot \xi$.)

11 Soient $\varphi : \mathbb{R}^N \to \mathbb{R}^M$, $\lambda : \mathbb{R}^N \to \mathbb{R}$, $f(x) = \lambda(x)\varphi(x)$; on suppose que $\varphi(0) = 0$, φ dérivable au point 0, λ continue en 0. Démontrer que f est dérivable au point 0 et $f'(0) = \lambda(0)\varphi'(0)$.

12 Vérifier que l'équation aux dérivées partielles

$$\frac{\partial u}{\partial y} = \frac{x}{x^2 + y^2}, \quad (x, y) \neq (0, 0)$$

possède comme solution générale

$$u(x, y) = \begin{cases} \operatorname{arc\,tg} \dfrac{y}{x} + \alpha(x) & \text{si } x \neq 0 \\ \alpha_1 & \text{si } x = 0,\ y > 0 \\ \alpha_2 & \text{si } x = 0,\ y < 0 \end{cases}$$

où $\alpha : \mathbb{R}^* \to \mathbb{R}$ est une fonction arbitraire et α_1, α_2 sont deux nombres réels quelconques. En déduire que l'équation ne possède pas de solution continue dans $\mathbb{R}^2 \setminus \{(0, 0)\}$.

13 Soit Ω un ensemble ouvert dans $\mathbb{R}^2$. L'objectif de cet exercice est de voir dans quelle mesure les seules solutions de l'équation $\mathrm{D}_2 f = 0$ dans Ω sont de la forme $f(x, y) = \alpha(x)$.

Supposons que Ω est tel que si (x, y_1), (x, y_2) sont dans Ω alors $(x, y) \in \Omega$, $y_1 \leqslant y \leqslant y_2$. (Cette propriété est vérifiée si Ω est convexe ; cette propriété n'entraîne pas que Ω est connexe.) Alors toutes les solutions f de l'équation $\mathrm{D}_2 f = 0$ dans Ω sont de la forme $f(x, y) = \alpha(x)$ où α est une fonction réelle quelconque définie dans P, la projection de Ω sur l'axe des x :

$$P = \{x \in \mathbb{R} : \exists\, y \in \mathbb{R} \text{ tel que } (x, y) \in \Omega\}.$$

Pour l'ensemble $\Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$ (qui ne satisfait pas la propriété proposée) la conclusion ci-dessus est en défaut.

14 Déterminer u telle que

$$\frac{\partial u}{\partial x} = \frac{x^2+y^2}{x}, \quad u(1,y) = \sin y \ (x > 0,\ y \in \mathbb{R}) .$$

Est-ce que u est unique en tant que fonction réelle définie sur $\{(x,y) : x \neq 0\}$?

15 Soit $\Omega = \mathbb{R}^2 \setminus \{(x,0) : x \geqslant 0\}$; démontrer qu'il existe $f \in C^1(\Omega)$ tel que $\mathrm{D}_2 f = 0$ dans Ω mais que $f(x,y)$ n'est pas de la forme $\alpha(x)$ (cf. Ex. 13).

16 Soit $f(x,y) = x^4 + x^3y + x^2y^2 + xy^3$. Vérifier que $(0,0)$ est un point critique de f qui n'est pas un extremum local pour f. (Noter que $f(x,y) = x(x+y)(x^2+y^2)$.)

17 Soit

$$f(x_1, \ldots, x_n) = (x_1 \ldots x_n) + \sum_{i=1}^{n} \frac{a_i}{x_i}$$

où les a_i et les x_i sont strictement positifs. Vérifier que l'unique point critique de f est $(\alpha_1, \ldots, \alpha_n)$ où

$$\alpha_i = a_i (a_1 \ldots a_n)^{-\frac{1}{n+1}}, \quad i = 1, \ldots, n .$$

Démontrer ensuite que l'infimum de f est atteint à ce point critique d'où, pour $x_1 > 0, \ldots, x_n > 0$,

$$(x_1 \ldots x_n) + \sum_{i=1}^{n} \frac{a_i}{x_i} \geqslant f(\alpha_1, \ldots, \alpha_n) = (n+1)(a_1 \ldots a_n)^{\frac{1}{n+1}}$$

avec égalité seulement si $x_i = \alpha_i$, $i = 1, \ldots, n$. En prenant $x_1 = x_2 = \ldots = x_n = 1$ on obtient

$$1 + \sum_{i=1}^{n} a_i \geqslant (n+1)(a_1 \ldots a_n)^{\frac{1}{n+1}} ;$$

d'où on déduit l'inégalité entre les moyennes arithmétiques et géométriques : si $a_1 > 0, \ldots a_N > 0$, on a

$$\frac{1}{N}(a_1 + \cdots + a_N) \geqslant (a_1 \ldots a_N)^{1/N}$$

avec égalité seulement si $a_1 = \ldots = a_N$.

18 Soit $\varphi(x) = -x \ln x$, $0 \leqslant x \leqslant 1$, avec $\varphi(0) = 0$; on vérifie que φ est continue, $\varphi(0) = \varphi(1) = 0$, φ' existe et est strictement monotone dans $]0, 1[$. Soit $H(p_1, \ldots, p_n) = \sum_{i=1}^{n} \varphi(p_i)$ où les p_i sont $\geqslant 0$ avec $\sum_{i=1}^{n} p_i = 1$; si

$$P = \left\{ (p_1, \ldots, p_n) : 0 \leqslant p_i \leqslant 1,\ i = 1, \ldots, n,\ \sum_{i=1}^{n} p_i = 1 \right\}$$

on établit par récurrence sur $n \geqslant 2$ que $\sup_{p \in P} H(p)$ est atteint pour $p \in \mathring{P}$. On vérifie ensuite que le seul point critique de f, fonction de $(n-1)$ variables $p_1, \ldots, p_{n-1}$, définie par

$$f(p_1, \ldots, p_{n-1}) = \sum_{i=1}^{n-1} \varphi(p_i) + \varphi(1 - p_1 - \ldots - p_{n-1}),$$

$$0 < p_i < 1, \quad i = 1, \ldots, n-1, \quad \sum_{i=1}^{n-1} p_i < 1,$$

est donné par $p_i = 1/n$, $i = 1, \ldots, n-1$. On conclut que

$$H(p_1, \ldots, p_n) \leqslant n\,\varphi\left(\frac{1}{n}\right) = \ln n$$

avec égalité seulement si $p_i = 1/n$, $i = 1, \ldots, n$.

19 Soit f une fonction dérivable définie dans un voisinage ouvert U de $a \in \mathbb{R}^N$ à valeurs dans $\mathbb{R}^N$; si $f(a) = b$ et g est une fonction définie dans un voisinage ouvert de b telle que $g\big(f(x)\big) = x$, $u \in U$, démontrer que g ne peut pas être dérivable au point b si $\det f'(a) = 0$.

20 Soit $f : \Omega \to \mathbb{R}^N$, $f = (f_1, \ldots, f_N)$, une fonction dérivable définie dans un ensemble ouvert convexe $\Omega \subset \mathbb{R}^N$. Supposons que

$$\sum_{i,j=1}^{N} \mathrm{D}_j f_i(x) t_i t_j > 0, \quad x \in \Omega.$$

pour tout choix de $(t_1, \ldots, t_N) \in \mathbb{R}^N \setminus \{(0, \ldots, 0)\}$. Démontrer que f est injective. (Si $a \neq b$, posons $g(x) = \{f(x) - f(a)\} \cdot (b - a)$; si $f(a) = f(b)$, on aurait $g(a) = g(b) = 0$ d'où l'existence d'un c tel que $\nabla g(c) \cdot (b - a) = 0$; un calcul explicite de cette dernière expression montre que la condition imposée sur de $\mathrm{D}_j f_i$ est alors violée.)

21 Soit

$$f(x, y) = \left(xy\,,\, -\frac{x^2}{2} + \theta(y)\right), \quad x \in \mathbb{R},\ y > 0,$$

$\theta'(y) > 0$ pour tout $y > 0$; en utilisant l'exercice 20, démontrer que f est injective dans $\{(x, y) : y > 0\}$.

CHAPITRE 3

Dérivées d'ordre supérieur

Les propositions principales de ce chapitre sont les suivantes : égalité des dérivée mixtes (sous conditions assez générales), formule de Taylor (surtout d'ordre 2). Plusieurs applications et exemples illustrent la grande utilité de ces propositions. A noter en particulier les applications à la recherche des extremums, aux fonctions harmoniques et aux fonctions convexes. On introduit une terminologie appropriée pour la notion de difféomorphisme entre deux ensembles ; c'est une précision de la notion de homéomorphisme déjà vue dans le chapitre 1.

3.1 DÉFINITIONS ET NOTATIONS PRINCIPALES

3.1.1 Généralités

Soit Ω une partie ouverte de $\mathbb{R}^N$. Dans le paragraphe 2.1.5 on a déjà introduit la classe $C^1(\Omega)$ de fonctions réelles continûment dérivables dans Ω ; il s'agit des fonctions $f : \Omega \to \mathbb{R}$ telles que les $\mathrm{D}_i f$ existent et sont continues dans Ω, $i = 1, \ldots, N$. Rappelons que, d'après la proposition 2.1.2(iii), une fonction f de $C^1(\Omega)$ est dérivable dans Ω. Les $\mathrm{D}_i f$, $i = 1, \ldots, N$, s'appellent les ***dérivées partielles de premier ordre*** de f ; ainsi $f \in C^1(\Omega)$ si ses dérivées partielles de premier ordre sont continues dans Ω.

On dit que f est ***deux fois continûment dérivable*** dans Ω (en symboles : $f \in C^2(\Omega)$) si $f \in C^1(\Omega)$ et toutes les dérivées partielles de premier ordre de f sont aussi de classe $C^1(\Omega)$; en symboles,

$$f \in C^2(\Omega) \quad \text{si } \mathrm{D}_i f \in C^1(\Omega),\ \mathrm{D}_j(\mathrm{D}_i f) \in C(\Omega),\ i \leqslant i, j \leqslant N\,.$$

Les quantités $\mathrm{D}_j(\mathrm{D}_i f)$ s'appellent les ***dérivées partielles de second ordre*** (ou ***d'ordre deux***) de f : si $i \neq j$, $\mathrm{D}_j(\mathrm{D}_i f)$ s'appelle une ***dérivée mixte d'ordre deux***. Un théorème important (sect. 3.2) dit que si $f \in C^2(\Omega)$, alors

$$\mathrm{D}_i(\mathrm{D}_j f) = \mathrm{D}_j(\mathrm{D}_i f)$$

de sorte que l'on peut écrire, sans ambiguïté, $\mathrm{D}_{ij} f$ pour ces dernières fonctions. Ainsi, pour $f \in C^2(\Omega)$, on aura $N + N(N-1)/2 = N(N+1)/2$ dérivées partielles d'ordre 2 (de type $\mathrm{D}_{ij} f$) au lieu de N^2. Si $f \in C^2(\Omega)$, nous dirons aussi que f est ***de classe*** C^2 dans Ω.

Pour un entier k positif quelconque, on définit par récurrence la notion de fonction réelle f ***de classe*** C^k dans Ω (en symboles : $f \in C^k(\Omega)$) ; on dit aussi que f est k ***fois continûment dérivable*** dans Ω. Pour $f \in C^k(\Omega)$ les ***dérivées partielles d'ordre*** k

$$\mathrm{D}_{i_1}(\mathrm{D}_{i_2} \ldots (\mathrm{D}_{i_k} f) \ldots), \quad 1 \leqslant i_1, \ldots, i_k \leqslant N$$

existent et sont des fonctions continues réelles dans Ω. On montre (sect. 3.2) que ces dérivées ne dépendent pas de la succession dans laquelle les dérivations sont faites de sorte que l'on peut écrire, sans ambiguïté, $\mathrm{D}_{i_1 \ldots i_k} f$ pour ces dérivées. On peut démontrer que pour $f \in C^k(\Omega)$, on aura (cf. Feller [19, p. 37])

$$\binom{N+k-1}{k} = \frac{N(N+1)\ldots(N+k-1)}{k!}$$

dérivées partielles d'ordre k. Si α_j est égal au nombre de fois que j apparaît dans la suite $i_1, \ldots, i_k$ et $\alpha = (\alpha_1, \ldots, \alpha_N)$ on écrit aussi

$$\mathrm{D}^\alpha f \text{ ou } \mathrm{D}_1^{\alpha_1} \ldots \mathrm{D}_N^{\alpha_N} f \quad \text{pour } \mathrm{D}_{i_1 \ldots i_k} f \, ;$$

ici, $\alpha_j \geqslant 0$, $1 \leqslant j \leqslant N$, $\alpha_1 + \cdots + \alpha_N = k$; si certains α_j sont nuls, on les omet en général dans l'écriture $\mathrm{D}_1^{\alpha_1} \ldots \mathrm{D}_N^{\alpha_N} f$.

Si $f : \Omega \to \mathbb{R}$ est une fonction telle que f est k fois continûment dérivable pour tout entier k positif, on dit que f est ***indéfiniment dérivable*** et on écrit $f \in C^\infty(\Omega)$; on dit aussi que f est ***de classe*** C^∞ dans Ω.

Si $f \in C^2(\Omega)$, la matrice symétrique de type (N, N),

$$[\mathrm{D}_{ij} f], \quad 1 \leqslant i, j \leqslant N,$$

s'appelle la ***matrice hessienne*** de f.

Pour avoir une notation cohérente, nous convenons qu'une fonction réelle f continue dans Ω est de classe C^0 ; nous écrivons aussi quelquefois $f \in C^0(\Omega)$ au lieu d'écrire $f \in C(\Omega)$, notation introduite dans le chapitre 1. Ainsi, on a

$$C^0(\Omega) \supset C^1(\Omega) \supset \ldots \supset C^\infty(\Omega) \, ;$$

il est facile de constater que ces inclusions sont strictes dès que Ω est une partie ouverte non vide de $\mathbb{R}^N$ (Ex. 18, § 3.2.6). Chaque $C^k(\Omega)$, $k \geqslant 0$, est un espace vectoriel réel et même une algèbre, c'est-à-dire, qu'en plus, si f et g sont dans $C^k(\Omega)$ alors leur produit $fg \in C^k(\Omega)$.

Si $f : \Omega \to \mathbb{R}^M$, $f = (f_1, \ldots, f_M)$, on dira que la fonction vectorielle f est ***de classe*** C^k dans Ω ou que f est k ***fois continûment dérivable*** si chaque composante $f_i \in C^k(\Omega)$, $i = 1, \ldots, M$; on écrira alors que $f \in C^k(\Omega\,; \mathbb{R}^M)$. De même, on dit que f est ***indéfiniment dérivable*** dans Ω ou que f est ***de classe*** C^∞ dans Ω si $f \in C^k(\Omega\,; \mathbb{R}^M)$ pour tout entier k positif ; on écrira $f \in C^\infty(\Omega\,; \mathbb{R}^M)$ pour la classe des fonctions indéfiniment dérivables à valeurs dans $\mathbb{R}^M$.

3.1.2 Remarques

Les notations en usage pour $D_{i_1 \dots i_k} f$ sont très diverses. Par exemple, pour $D_{ij} f$ on trouvera dans la littérature scientifique les notations suivantes :

$$\frac{\partial^2 f}{\partial x_i \partial x_j}, \quad f''_{x_i x_j}, \quad f_{x_i x_j}, \quad f''_{ij}, \dots$$

L'écriture $D^\alpha f$ ou $D_1^{\alpha_1} \dots D_N^{\alpha_N} f$ est souvent commode ; avec la notation $D^\alpha f$ certaines formules (comme celle de Taylor donnée dans la suite) deviennent simples à exprimer. Heureusement, dans un grand nombre d'applications seules les dérivées d'ordre un et deux jouent un rôle important, et pour celles-ci nous utiliserons généralement la notation $D_i f$, $D_{ij} f$.

Pour illustrer les notations, prenons quelques exemples. Soit

$$f(x, y) = x^2 \sin xy$$

où x, y sont deux variables réelles. Alors

$$D_1\big(f(x,y)\big) = 2x \sin xy + x^2 y \cos xy, \quad D_2 f(x,y) = x^3 \cos xy$$

$$\begin{aligned} D_2(D_1 f)(x,y) &= 2x^2 \cos xy + x^2 \cos xy - x^3 y \sin xy \\ &= 3x^2 \cos xy - x^3 y \sin xy = D_1(D_2 f)(x,y) = D_{12} f(x,y) \end{aligned}$$

où nous avons identifié x comme la variable numéro 1 et y comme la variable numéro 2. On a ici une vérification de l'égalité

$$D_2 D_1 f = D_1 D_2 f$$

qui justifie l'écriture $D_{12} f$ pour la dérivée mixte en question ; évidemment, on pourra aussi écrire $D_{21} f$ pour la même dérivée ou encore

$$\frac{\partial^2 f}{\partial x \partial y}, \quad f''_{xy}, \dots$$

Soit $f(x) = x_1 x_2^2 x_3^3$ où $x = (x_1, x_2, x_3) \in \mathbb{R}^3$. Alors,

$$D_1^2 f(x) = 0, \quad D_1 D_2^2 f(x) = 2x_3^3, \quad D_2^2 D_3^3 f(x) = 12\, x_1 ;$$

$$\begin{aligned} D_1^2 f = D_{11} f = D^\alpha f, \quad \alpha = (2,0,0) \\ D_1 D_2^2 f = D_{122} f = D^\alpha f, \quad \alpha = (1,2,0) \\ D_2^2 D_3^3 f = D_{22\,333} f = D^\alpha f, \quad \alpha = (0,2,3). \end{aligned}$$

3.1.3 Certains opérateurs différentiels

Certaines combinaisons de dérivées partielles d'une fonction apparaissent fréquemment dans les considérations les plus diverses. Il est utile d'introduire la terminologie, ***opérateur différentiel***, pour ces expressions ; nous en donnons quelques exemples importants.

Soit Ω une partie ouverte de $\mathbb{R}^N$; une expression du type

$$Tf(x) = \sum_{i=1}^{N} a_i(x)\, \mathrm{D}_i f(x) + a_0(x) f(x)\,, \tag{3.1}$$

où les a_i sont des fonctions réelles définies dans Ω et $f \in C^1(\Omega)$, s'appelle un ***opérateur différentiel linéaire du premier ordre***. Notons que

$$T(\alpha f + \beta g) = \alpha T f + \beta T g$$

pour tout choix de α, β dans $\mathbb{R}$ et de f, g dans $C^1(\Omega)$; si les $a_i \in C^0(\Omega)$, $i = 0, 1, \ldots, N$, alors (3.1) donne une application linéaire

$$T : C^1(\Omega) \to C^0(\Omega)\,.$$

De même, une expression du type

$$Tf(x) = \sum_{1 \leqslant i,\, j \leqslant N} a_{ij}(x) \mathrm{D}_{ij} f(x) + \sum_{1 \leqslant i \leqslant N} a_i(x) \mathrm{D}_i f(x) + a_0(x) f(x)\,, \tag{3.2}$$

où les coefficients a_{ij}, a_i sont des fonctions réelles définies dans Ω et $f \in C^2(\Omega)$, s'appelle un ***opérateur différentiel linéaire du second ordre***. Ici aussi, on voit immédiatement que pour T donné par (3.2), on a

$$T(\alpha f + \beta g) = \alpha T f + \beta T g$$

pour α, β dans $\mathbb{R}$ et f, g dans $C^2(\Omega)$; si les a_{ij}, a_i sont continues, alors (3.2) définit une application linéaire

$$T : C^2(\Omega) \longrightarrow C^0(\Omega)\,.$$

Nous notons quelques exemples importants du type (3.2) (avec coefficients constants) :

- l'***opérateur de Laplace***

$$\triangle f(x) = \sum_{i=1}^{N} \mathrm{D}_{ii} f(x)\ ; \tag{3.3}$$

- l'***opérateur de l'équation de la chaleur***

$$Tf(t,x) = \mathrm{D}_t f(t,x) - k\, \triangle f(t,x)\ ; \tag{3.4}$$

  ici k est une constante > 0, $t \in \mathbb{R}$, $x \in \mathbb{R}^N$;
- l'***opérateur de l'équation des ondes*** ou l'***opérateur de D'Alembert***

$$\Box f(t,x) = \mathrm{D}_{tt} f(t,x) - c^2\, \triangle f(t,x)\ ; \tag{3.5}$$

  ici c est une constante > 0, $t \in \mathbb{R}$, $x \in \mathbb{R}^N$.

Dans (3.4) et (3.5) on suppose que f est fonction réelle définie dans un ensemble ouvert de $\mathbb{R}^{N+1}$ dont les points sont notés (t,x), $t \in \mathbb{R}$, $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ et que l'opérateur Δ opère sur f comme fonction de x avec t fixe et D_t, D_{tt} opèrent sur f comme fonction de t avec x fixe ; ainsi

$$\Box f(t,x) = \frac{\partial^2 f}{\partial t^2}(t,x) - c^2 \sum_{i=1}^{N} \frac{\partial^2 f}{\partial x_i^2}(t,x).$$

Une fonction $f \in C^2(\Omega)$, Ω étant une partie ouverte de $\mathbb{R}^N$, s'appelle une ***fonction harmonique*** dans Ω si

$$\Delta f(x) = 0, \quad x \in \Omega. \tag{3.6}$$

Les fonctions harmoniques jouent un rôle important dans plusieurs théories mathématiques et physiques. Certains auteurs écrivent ∇^2 pour Δ.

De ce qui précède, on voit immédiatement comment définir des opérateurs différentiels linéaires d'ordre supérieur. Par exemple

$$\Delta^2 f(x) = \Delta(\Delta f)(x)$$

définit un opérateur différentiel linéaire d'ordre 4 (à coefficients constants) ; pour $x = (x_1, x_2)$, cet opérateur peut s'écrire plus explicitement comme suit :

$$\Delta^2 f(x) = \Delta(\mathrm{D}_{11} f + \mathrm{D}_{22} f)(x) = \mathrm{D}_1^4 f(x) + 2\mathrm{D}_1^2 \mathrm{D}_2^2 f(x) + \mathrm{D}_2^4 f(x).$$

Ici, f est de classe C^4 de sorte qu'une dérivée partielle d'ordre 4 du type $\mathrm{D}_{11}\mathrm{D}_{22} f$ peut s'écrire comme $\mathrm{D}_1^2 \mathrm{D}_2^2 f$, l'ordre successif des dérivations étant sans importance. Les fonctions f telles que

$$\Delta^2 f(x) = 0, \quad x \in \Omega \subset \mathbb{R}^N, \tag{3.7}$$

s'appellent les ***fonctions biharmoniques*** dans Ω, une partie ouverte de $\mathbb{R}^N$; pour $N = 2$, ces fonctions et l'opérateur Δ^2 jouent un rôle important dans la théorie d'élasticité.

Les opérateurs différentiels non linéaires deviennent de plus en plus importants dans les travaux récents ; leur théorie est difficile et, en grande partie, incomplète. Mais, on les trouve dans presque toutes les théories physiques qui vont un peu au-delà des approximations dites linéaires. Par exemple, l'opérateur

$$Tf = (\mathrm{D}_{11} f)^2 + (\mathrm{D}_{11} f)(\mathrm{D}_{22} f) - 2(\mathrm{D}_{12} f)^2 \tag{3.8}$$

où f est une fonction de classe C^2 de $x = (x_1, x_2)$ dans une partie ouverte de $\mathbb{R}^2$ apparaît aussi dans la théorie d'élasticité bidimensionnelle ; il est clair que T est non linéaire.

3.2 DÉRIVÉES MIXTES

Dans cette section nous allons établir que si une fonction réelle f de N variables réelles $x_1, \ldots, x_N$ est k fois continûment dérivable, alors ses dérivées partielles d'ordre k, $\mathrm{D}_{i_1}\mathrm{D}_{i_2}\ldots\mathrm{D}_{i_k}f$, peuvent être calculées en dérivant f successivement par rapport aux variables $x_{i_1}, \ldots, x_{i_k}$ dans un ordre quelconque. On exprime ceci en disant que l'expression $\mathrm{D}_{i_1}\ldots\mathrm{D}_{i_k}f$ est invariante sur toute permutation des indices $i_1, \ldots, i_k$ — pourvu que f soit de classe C^k. Ce fait justifie l'introduction de la notation $\mathrm{D}^\alpha f$ dans la section 3.1. Même s'il existe des exemples où les dérivées $\mathrm{D}_i\mathrm{D}_j f$ et $\mathrm{D}_j\mathrm{D}_i f$ sont inégales en certains points, dans la plupart des applications des fonctions de plusieurs variables, on travaille sous des hypothèses qui rendent les dérivées partielles indépendantes de l'ordre dans lequel les dérivations sont faites.

3.2.1 Proposition

Soit f une fonction réelle deux fois continûment dérivable dans une partie ouverte Ω de $\mathbb{R}^N$. Alors, pour $i \neq j$, $1 \leqslant i, j \leqslant N$,

$$\mathrm{D}_i(\mathrm{D}_j f)(x) = \mathrm{D}_j(\mathrm{D}_i f)(x), \quad x \in \Omega.$$

Nous déduisons cette proposition d'une autre plus générale qui n'est pas plus difficile à démontrer.

3.2.2 Proposition (*théorème de Schwarz*)

Soit $f : \Omega \to \mathbb{R}$, Ω une partie ouverte de $\mathbb{R}^N$. Supposons que $\mathrm{D}_i f$, $\mathrm{D}_j f$, $\mathrm{D}_i(\mathrm{D}_j f)$ existent dans Ω et que $\mathrm{D}_i(\mathrm{D}_j f)$ est continue en un point $\xi \in \Omega$; alors $\mathrm{D}_j(\mathrm{D}_i f)(\xi)$ existe et

$$\mathrm{D}_i(\mathrm{D}_j f)(\xi) = \mathrm{D}_j(\mathrm{D}_i f)(\xi).$$

DÉMONSTRATION

Pour simplifier l'écriture, nous considérons le cas d'une fonction f de deux variables réelles, notées (x, y) ; comme la proposition ne concerne que la dérivation par rapport à deux des N variables, les autres étant tenues fixes, il suffit de démontrer celle-ci dans le cas $N = 2$. Aussi, nous écrivons $\mathrm{D}_1 f$ pour la dérivée partielle par rapport à x, $\mathrm{D}_2 f$ pour la dérivée partielle par rapport à y, $\mathrm{D}_1\mathrm{D}_2 f$ au lieu de $\mathrm{D}_1(\mathrm{D}_2 f)$ et $\mathrm{D}_2\mathrm{D}_1 f$ pour $\mathrm{D}_2(\mathrm{D}_1 f)$; nous considérons un point $\xi = (a, b) \in \Omega \subset \mathbb{R}^2$ où $\mathrm{D}_1\mathrm{D}_2 f$ est supposé être continue. Sous l'hypothèse que $\mathrm{D}_1 f$, $\mathrm{D}_2 f$, $\mathrm{D}_2\mathrm{D}_1 f$ existent dans Ω, nous démontrons que $\mathrm{D}_1\mathrm{D}_2 f(\xi)$ existe et est égal à $\mathrm{D}_2\mathrm{D}_1 f(\xi)$.

Posons, pour s, t tels que $(a + s, b + t) \in \Omega$,

$$E(s,t) = f(a+s, b+t) - f(a+s, b) - f(a, b+t) + f(a, b) ;$$

la démonstration se fait en calculant $E(s,t)/st$ de deux façons différentes. Soit

$$\varphi(x) = f(x, b+t) - f(x,b)\ ;$$

alors

$$\varphi'(x) = \mathrm{D}_1 f(x, b+t) - \mathrm{D}_1 f(x,b)$$

et

$$E(s,t) = \varphi(a+s) - \varphi(a) = s\varphi'(a + c_1 s)$$

où $0 < c_1 < 1$; on en déduit que

$$E(s,t) = s\big\{\mathrm{D}_1 f(a + c_1 s, b+t) - \mathrm{D}_1 f(a + c_1 s, b)\big\} = st\mathrm{D}_2\mathrm{D}_1 f(a + c_1 s, b + c_2 t)$$

où $0 < c_2 < 1$.

La continuité de $\mathrm{D}_2\mathrm{D}_1 f$ au point (a,b) donne

$$\lim_{(s,t)\to(0,0)} \frac{E(s,t)}{st} = \mathrm{D}_2\mathrm{D}_1 f(a,b)$$

ce qui équivaut à l'affirmation suivante : pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que

$$\left|\frac{E(s,t)}{st} - \mathrm{D}_2\mathrm{D}_1 f(a,b)\right| \leqslant \varepsilon \quad \text{si } 0 < |s| < \delta\,,\ 0 < |t| < \delta\,. \tag{3.9}$$

D'autre part, on peut aussi écrire

$$\varphi(x) = t\big\{\mathrm{D}_2 f(x,b) + \rho(x,t)\big\}$$

où

$$\lim_{t\to 0} \rho(x,t) = 0$$

pour chaque x fixé. Ainsi,

$$E(s,t) = \varphi(a+s) - \varphi(a) = t\big\{\mathrm{D}_2 f(a+s,b) - \mathrm{D}_2 f(a,b) + \rho(a+s,t) - \rho(a,t)\big\}$$

d'où

$$\frac{E(s,t)}{st} = \frac{1}{s}\left\{\mathrm{D}_2 f(a+s,b) - \mathrm{D}_2 f(a,b)\right\} + \frac{1}{s}\left\{\rho(a+s,t) - \rho(a,t)\right\}$$

ce qui donne

$$\lim_{t\to 0} \frac{E(s,t)}{st} = \frac{1}{s}\left\{\mathrm{D}_2 f(a+s,b) - \mathrm{D}_2 f(a,b)\right\} \tag{3.10}$$

pour tout s fixé, en particulier tels que $0 < |s| < \delta$. En combinant (3.9) et (3.10) on voit que, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que

$$\left|\frac{1}{s}\left\{\mathrm{D}_2 f(a+s,b) - \mathrm{D}_2 f(a,b)\right\} - \mathrm{D}_2\mathrm{D}_1 f(a,b)\right| \leqslant \varepsilon$$

dès que $0 < |s| < \delta$; ceci établit l'existence de $\mathrm{D}_1\mathrm{D}_2 f(a,b)$ ainsi que son égalité avec $\mathrm{D}_2\mathrm{D}_1 f(a,b)$. ♦

3.2.3 Remarque

La dernière partie de la démonstration précédente consiste à établir le *lemme* suivant (où l'on prendra $g(s,t) = E(s,t)/st$) : soit $g : V \setminus \{(0,0)\} \to \mathbb{R}$ où V est un voisinage ouvert de $(0,0)$; si

(i) $\lim_{(s,t)\to(0,0)} g(s,t) = A$ existe,
(ii) $\lim_{t\to 0} g(s,t) = A(s)$ existe pour chaque s fixé, alors

$$\lim_{s\to 0} A(s) = A\,.$$

En prenant $g(s,t) = s\sin(1/t)$ ou (i) est vérifiée mais (ii) est en défaut, on voit l'importance d'une hypothèse du type (ii).

3.2.4 Proposition

Si $f \in C^k(\Omega)$, Ω une partie ouverte de $\mathbb{R}^N$, k un entier $\geqslant 2$, alors

$$\mathrm{D}_{i_1}\mathrm{D}_{i_2}\ldots\mathrm{D}_{i_k}f(x) = \mathrm{D}_{j_1}\mathrm{D}_{j_2}\ldots\mathrm{D}_{j_k}f(x)\,,\quad x\in\Omega$$

où $(j_1, \ldots, j_k)$ est une permutation de $(i_1, \ldots, i_k)$.

La démonstration de cette proposition est immédiate à partir de la proposition 3.2.1 ; elle se fait par récurrence. Au lieu d'écrire une démonstration formelle, nous considérons le cas $k = 3$ et démontrons, à titre d'exemple que

$$\mathrm{D}_{i_1}\,\mathrm{D}_{i_2}\,\mathrm{D}_{i_3}f = \mathrm{D}_{i_2}\,\mathrm{D}_{i_1}\,\mathrm{D}_{i_3}f$$

où i_1, i_2, i_3 sont trois indices quelconques choisis parmi $\{1, 2, \ldots, N\}$. En effet, cette égalité est banale si $i_1 = i_2$; si $i_1 \neq i_2$, nous considérons la fonction $g = \mathrm{D}_{i_3}f$; si f est de classe C^3, g sera de classe C^2 de sorte que la proposition 3.2.1 appliquée à g nous donnera l'égalité en question.

3.2.5 Remarques et exemples

La proposition 3.2.2 est assez générale pour garantir l'égalité des dérivées mixtes pour la plupart des fonctions rencontrées. Noter que la continuité des dérivées partielles $\mathrm{D}_i f$, $\mathrm{D}_j f$ n'est pas exigée dans cette proposition ; seule la continuité de $\mathrm{D}_i\mathrm{D}_j f$ garantit l'existence de $\mathrm{D}_j\mathrm{D}_i f$ et leur égalité. Si f est une fonction réelle des variables réelles x, y de la forme

$$f(x,y) = a(x) + b(y)\,,$$

on voit que

$$\mathrm{D}_1 f(x,y) = a'(x)\,,\quad D_2 f(x,y) = b'(y)$$
$$\mathrm{D}_2\mathrm{D}_1 f(x,y) = 0 = \mathrm{D}_1\mathrm{D}_2 f(x,y)$$

dès que a, b sont dérivables. Ceci donne des exemples des f pour lesquels $D_1 f$, $D_2 f$ ne sont pas nécessairement continues mais où les dérivées mixtes sont continues et égales.

Si f est de la forme

$$f(x,y) = a(x)b(y)$$

où a, b sont des fonctions réelles dérivables, x, y étant des variables réelles, on aura

$$D_1 D_2 f(x,y) = a'(x)b'(y) = D_2 D_1 f(x,y)$$

sans que les dérivées mixtes soient nécessairement continues.

Nous donnons maintenant une famille d'exemples simples où les dérivées mixtes ne sont pas égales partout. Soit

$$f(x,y) = \begin{cases} xy\,h(x,y) & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases} \tag{3.11}$$

où $h \in C^2\big(\mathbb{R}^2 \setminus \{(0,0)\}\big)$, h étant bornée dans un voisinage ouvert de $(0,0)$ et telle que

$$\lim_{x\to 0} h(x,0) = \alpha\,, \quad \lim_{y\to 0} h(0,y) = \beta\,, \quad \alpha \neq \beta\,.$$

Evidemment, $\lim_{(x,y)\to(0,0)} h(x,y)$ n'existe pas ; néanmoins, f est continue au point $(0,0)$. Aussi, $D_1 f$, $D_2 f$, $D_1 D_2 f$, $D_2 D_1 f$ existent partout mais

$$D_1 D_2 f(0,0) = \alpha\,, \quad D_2 D_1 f(0,0) = \beta$$

(Ex. 14-15, § 3.2.6).

3.2.6 Exercices

1 Calculer les dérivées partielles du second ordre des fonctions données par les expressions suivantes :

(a) $\cos xy$;
(b) e^{-xy} ;
(c) $\sin(x+2y)$;
(d) $4xy + y^2 + 2x$;
(e) $\exp\sin(x-y)$;
(f) $x\exp(x^2+y^2)$;
(g) $\operatorname{arc\,tg}(x-y)$;
(h) $x^2 y + 2\sin(xy)$.

Vérifier chaque fois l'égalité des dérivées mixtes.

2 Calculer les dérivées partielles d'ordre $n \geqslant 2$ des fonctions données par les expressions suivantes :

(a) xe^y ; (b) y^2/x ; (c) $(xy)^{-1}$.

3 Si $u(t,x) = f(x+at) + g(x-at)$ démontrer que

$$u''_{tt} = a^2 u''_{xx} \quad \text{(équation des ondes)}$$

où a est une constante > 0 et f, g sont deux fonctions réelles quelconques d'une variable réelle, deux fois dérivables.

4 Soit $r(x) = \|x\|$, $x = (x_1, \ldots, x_N) \in \mathbb{R}^N \setminus \{0\}$. Vérifier successivement que

$$r(x)\,\mathrm{D}_j r(x) = x_j\,, \quad \sum_{j=1}^{N} \bigl(\mathrm{D}_j r\bigr)^2 = 1\,, \quad \Delta r = \frac{N-1}{r}\,.$$

5 Soit $f = h \circ r$ où r est comme dans l'exercice 4 et $h : \,]0, \infty[\,\to \mathbb{R}$ est deux fois dérivable. Vérifier que

$$\Delta f = h'' \circ r + \frac{N-1}{r}\, h' \circ r\,.$$

En déduire que si $f = r^{-N+2}$ et $N \geqslant 3$, alors f est harmonique dans $\mathbb{R}^N \setminus \{0\}$; démontrer la même affirmation pour $N = 2$ et $f = \ln r$.

6 Vérifier que les fonctions définies par les expressions suivantes sont harmoniques :

(i) $x^3 - 3xy^2$;
(ii) $\mathrm{e}^x \cos y$;
(iii) $x/(x^2+y^2)$ dans $\mathbb{R}^2 \setminus \{(0,0)\}$;
(iv) $\operatorname{arc\,tg}(y/x)$ dans $\{(x,y) : x > 0\}$.

7 Vérifier que toute fonction harmonique (de classe C^4 — en fait, elle est automatiquement de classe C^∞) est biharmonique.

8 Vérifier la formule

$$\Delta(fg) = f\,\Delta g + g\,\Delta f + 2\nabla f \cdot \nabla g$$

pour deux fonctions f, g de N variables de classe C^2. En déduire que si f, g sont harmoniques dans un ensemble ouvert $\Omega \subset \mathbb{R}^N$, alors fg est aussi harmonique dans Ω si et seulement si $\nabla f \cdot \nabla g = 0$ dans Ω.

9 Si f est une fonction harmonique des variables $x_1, \ldots, x_N$ vérifier que $x_1 f(x_1, \ldots, x_N)$ donne une fonction biharmonique des mêmes variables.

10 Soit $u(t,x) = (\beta t)^\gamma \exp(-\alpha r^2/t)$ où $r = \|x\|$, $x \in \mathbb{R}^N$ et α, β, γ sont des constantes > 0 et $t > 0$. Vérifier que

$$\frac{\partial u}{\partial t} = k^2 \, \Delta u \quad \text{(équation de diffusion)}$$

où k est une constante > 0 si et seulement si

$$\alpha = \frac{1}{4k^2}, \quad \gamma = -\frac{N}{2}, \quad \beta \text{ quelconque}.$$

(Ici $\Delta u = \sum_{i=1}^N \mathrm{D}_{ii} u$, $\mathrm{D}_{ii} u = \partial^2 u / \partial x_i^2$, $x = (x_1, \ldots, x_N)$.)

11 Soit

$$u(t,x,y,z) = \frac{1}{r} f(r + ct), \quad r = \left(x^2 + y^2 + z^2\right)^{1/2} > 0 ;$$

Vérifier que

$$\frac{\partial^2 u}{\partial t^2} = c^2 \, \Delta u \quad \text{(équation des ondes)}$$

si f est de classe C^2 et Δ est appliqué sur u en tant que fonction de x, y, z avec t fixé.

12 Démontrer que l'opérateur défini dans (3.8) est non linéaire.

13 Démontrer le lemme énoncé dans le paragraphe 3.2.3.

14 Pour f de (3.11) vérifier que

$$\mathrm{D}_1 f(0,0) = \mathrm{D}_2 f(0,0) = 0, \quad \mathrm{D}_1 \mathrm{D}_2 f(0,0) = \alpha, \quad \mathrm{D}_2 \mathrm{D}_1 f(0,0) = \beta .$$

Voir que $h(x,y) = (x^2 + \lambda y^2)/(x^2 + y^2)$ si $(x,y) \neq (0,0)$, λ nombre réel $\neq 1$, est un choix possible pour h ; en fait, pour ce choix de h, $f \in C^1(\mathbb{R}^2)$.

15 Pour f de (3.11), vérifier que

$$E(s,t) = st \, h(s,t)$$

où $E(s,t)$ est défini au point $(0,0)$ comme dans la démonstration de la proposition 3.2.2 ; ceci montre que $\lim_{(s,t)\to(0,0)} \big(E(s,t)\big)/st$ n'existe pas dans ce cas.

16 Si $\mathrm{D}_1 f(0,0)$, $\mathrm{D}_2 f(0,0)$ existent et $E(s,t) \to 0$ lorsque $(s,t) \to (0,0)$ où $E(s,t) = f(s,t) - f(s,0) - f(0,t) + f(0,0)$, démontrer que f est continue au point $(0,0)$. En déduire que si f est une fonction de deux variables telle que $\mathrm{D}_1 f$, $\mathrm{D}_2 f$, $\mathrm{D}_2 \mathrm{D}_1 f$ existent dans un voisinage ouvert de $(0,0)$ et $\mathrm{D}_2 \mathrm{D}_1 f$ est continue au point $(0,0)$, alors f est continue à $(0,0)$.

17 Soit f une fonction de deux variables définie dans un voisinage ouvert de $(0,0)$. Si $s \mapsto f(s,0)$ est continue à $s=0$ et $t \mapsto f(0,t)$ est continue à $t=0$, alors f est continue à $(0,0)$ si et seulement si $E(s,t) \to 0$ lorsque $(s,t) \to (0,0)$ où $E(s,t)$ est comme dans l'exercice 16.

18 Soit Ω une partie ouverte de $\mathbb{R}^N$ et $k \in \mathbb{N}^*$. Démontrer qu'il existe une fonction $f \in C^{k-1}(\Omega)$ qui n'est pas dans $C^k(\Omega)$. (Il suffit de démontrer ceci lorsque $N=1$ et Ω est un intervalle ouvert de $\mathbb{R}$, par exemple $]-1,1[$; procéder par induction sur k dans ce cas.)

3.3 FORMULE DE TAYLOR

La dérivabilité d'une fonction f équivaut à la possibilité d'un certain type d'approximation locale de f par une fonction affine. On sait que dans le cas d'une fonction réelle f d'une variable réelle, si f possède n dérivées alors f peut être approchée localement par un polynôme de degré n, le degré d'approximation s'améliorant localement avec l'augmentation de n. Une formulation précise de cette idée est fournie par la formule de Taylor que nous généraliserons dans cette section au cas d'une fonction réelle de N variables réelles. La formule de Taylor s'appelle aussi le ***développement de Taylor*** ou le ***développement limité***. La généralisation au cas de N variables réelles sera obtenue en partant de la formule pour le cas d'une variable réelle que nous donnons ici dans la forme dans laquelle elle sera utilisée ici : soit φ une fonction réelle de classe C^n dans un intervalle ouvert de $\mathbb{R}$ contenant $[0,1]$; alors

$$\varphi(1) = \varphi(0) + \frac{1}{1!}\varphi'(0) + \frac{1}{2!}\varphi''(0) + \cdots + \frac{1}{(n-1)!}\varphi^{(n-1)}(0) + R_n \tag{3.12}$$

$$R_n = \int_0^1 \frac{(1-t)^{n-1}}{(n-1)!}\varphi^{(n)}(t)\,\mathrm{d}t = \frac{1}{n!}\varphi^{(n)}(\theta)$$

où θ est un nombre approprié dans $]0,1[$.

Rappelons que la démonstration de (3.12) se base sur une suite d'intégrations par parties opérées sur la formule intégrale pour R_n ; l'égalité entre les deux formules données pour R_n est une conséquence du théorème de la moyenne du calcul intégral élémentaire appliqué à la fonction continue $\varphi^{(n)}$.

3.3.1 Notation

Pour simplifier l'écriture des formules, il est utile d'introduire la notation des ***multi-indices***. Soit f une fonction réelle des N variables réelles

$x_1, \ldots, x_N$; un ***multi-indice*** à N variables est un N-tuple $\alpha = (\alpha_1, \ldots, \alpha_N)$ où chaque $\alpha_j \in \mathbf{N}$ et on pose

$$\mathrm{D}^\alpha f(x) = \mathrm{D}_1^{\alpha_1} \ldots \mathrm{D}_N^{\alpha_N} f(x), \quad x = (x_1, \ldots, x_N)$$

$$|\alpha| = \alpha_1 + \cdots + \alpha_N, \quad \alpha! = \alpha_1! \cdots \alpha_N!, \quad x^N = x_1^{\alpha_1} \ldots x_N^{\alpha_n} \quad (0! = 1).$$

Si $|\alpha| = k$ on dit que $\mathrm{D}^\alpha f(x)$ est une dérivée partielle de f d'ordre k évaluée au point x. Comme on l'a vu dans la section précédente, cette écriture est justifiée, si f est de classe C^k, dans le sens que l'ordre dans lequel les dérivations sont effectuées n'a pas d'influence. Avec cette notation, on a les formules suivantes :

- si f est de classe C^1,

$$\sum_{|\alpha|=1} \frac{(x-a)^\alpha}{\alpha!} \mathrm{D}^\alpha f(a) = \sum_{i=1}^N (x_i - a_i)\, \mathrm{D}_i f(a) = \nabla f(a) \cdot (x-a) \tag{3.13}$$

- si f est de classe C^2,

$$\begin{aligned} &\sum_{|\alpha|=2} \frac{(x-a)^\alpha}{\alpha!} \mathrm{D}^\alpha f(a) \\ &= \sum_{i=1}^N \frac{(x_i - a_i)^2}{2!} \mathrm{D}_{ii} f(a) + \sum_{1 \leqslant i < j \leqslant N} (x_i - a_i)(x_j - a_j) \mathrm{D}_{ij} f(a) \\ &= \frac{1}{2!} \sum_{1 \leqslant i,\, j \leqslant N} (x_i - a_i)(x_j - a_j) \mathrm{D}_{ij} f(a) = \frac{1}{2!} \big\{ H(a)(x-a) \big\} \cdot (x-a) \end{aligned} \tag{3.14}$$

où $H(a)$ est la matrice hessienne formée des secondes dérivées de f : $H(a) = \big[\mathrm{D}_{ij} f(a)\big]$; noter que $H(a)$ est une matrice symétrique parce que f est de classe C^2.

Nous laissons la vérification des formules (3.13) et (3.14) comme un exercice utile pour la compréhension de la notation des multi-indices. Plus généralement, si f est de classe C^k, on a

$$\sum_{|\alpha|=k} \frac{(x-a)^\alpha}{\alpha!} \mathrm{D}^\alpha f(a) = \frac{1}{k!} \sum_{1 \leqslant i_1, \ldots, i_k \leqslant N} (x_{i_1} - a_{i_1}) \cdots (x_{i_k} - a_{i_k}) \mathrm{D}_{i_1 \ldots i_k} f(a). \tag{3.15}$$

Pour établir (3.15), il suffit de calculer le nombre de k-tuples $(i_1, \ldots, i_k)$ (chaque i_j étant un entier de 1 à N) tels que m apparaît α_m fois dans ce k-tuple, $m = 1, \ldots, N$ et $\alpha_1 + \cdots + \alpha_m = k$; ce nombre est égal à

$$\binom{k}{\alpha_1} \binom{k-\alpha_1}{\alpha_2} \binom{k-\alpha_1-\alpha_2}{\alpha_3} \cdots = \frac{k!}{\alpha_1! \cdots \alpha_N!} = \frac{k!}{\alpha!}$$

où

$$\binom{k}{r} = \frac{k!}{r!\,(k-r)!}.$$

En effet,

$$\binom{k}{\alpha_1}$$

est le nombre de façons différentes d'insérer 1, α_1 fois, parmi les k places $i_1, \ldots, i_k$, à la suite de quoi il nous reste $(k - \alpha_1)$ places à remplir α_2 fois avec 2, etc. Ce raisonnement nous donne le nombre indiqué $k!/\alpha!$ qui s'appelle aussi un ***coefficient multinomial***. Ces coefficients apparaissent également dans l'identité dite ***multinomiale*** suivante

$$\begin{aligned}(X_1 + \cdots + X_N)^k &= \sum_{1 \leqslant i_1, \ldots, i_k \leqslant N} X_{i_1} X_{i_2} \cdots X_{i_N} \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha_1! \cdots \alpha_N!} X_1^{\alpha_1} \cdots X_N^{\alpha_N} = \sum_{|\alpha|=k} \frac{k!}{\alpha!} X^\alpha\end{aligned} \tag{3.16}$$

où $\alpha = (\alpha_1, \ldots, \alpha_N)$ et où $X_1, \ldots, X_N$ sont N symboles quelconques qui peuvent être multipliés et ajoutés suivant les lois de commutativité, associativité et distributivité usuelles ; autrement dit, cette identité est valable pour N éléments $X_1, \ldots, X_N$ d'un anneau commutatif quelconque. La formule (3.15) peut être considérée comme un cas particulier de l'identité multinomiale (3.16) écrite pour

$$(h_1 \mathrm{D}_1 + \cdots + h_N \mathrm{D}_N)^k$$

avec $h_i = (x_i - a_i)$ où les symboles $h_i \mathrm{D}_i$ sont considérés comme N éléments de l'anneau commutatif des ***opérateurs différentiels à coefficients constants*** (Ex. 7, § 3.3.5). Une notation assez répandue anciennement est d'écrire

$$h_1 \mathrm{D}_1 + \cdots + h_N \mathrm{D}_N = h \cdot \nabla\,, \quad h = (h_1, \ldots, h_N)\,,\ \nabla = (\mathrm{D}_i, \ldots, \mathrm{D}_N)$$

ce qui donne l'écriture suivante

$$(h \cdot \nabla)^k f(a) = \sum_{1 \leqslant i_1, \ldots, i_k \leqslant N} h_{i_1} \cdots h_{i_k} \mathrm{D}_{i_1} \ldots \mathrm{D}_{i_k} f(a)\,.$$

3.3.2 Proposition (*formule de Taylor : développement limité d'ordre n*)

Soit $f \in C^n(\Omega)$ où Ω est une partie ouverte de $\mathbb{R}^N$ et n un entier $\geqslant 1$; si a, x sont deux points de Ω tels que le segment $[a, x]$ est contenu dans Ω alors, en écrivant $x = (x_1, \ldots, x_N)$, $a = (a_1, \ldots, a_N)$, on a

$$f(x) = f(a) + \sum_{1 \leqslant |\alpha| < n} \frac{(x-a)^\alpha}{\alpha!} \mathrm{D}^\alpha f(a) + R_n(x; a) \tag{3.17}$$

où

$$R_n(x; a) = \sum_{|\alpha|=n} \frac{(x-a)^\alpha}{\alpha!} \mathrm{D}^\alpha f\big(a + \theta(x-a)\big) = \sum_{|\alpha|=n} \frac{(x-a)^\alpha}{\alpha!} \mathrm{D}^\alpha f(a) + \rho_n(x; a),$$

θ étant un nombre, dépendant de a et x, dans l'intervalle $]0, 1[$ *et*

$$|\rho_n(x;a)| \leqslant \frac{\delta(x)}{n!}\left(|x_1 - a_1| + \cdots + |x_N - a_N|\right)^n$$

$\delta(x) = \sup\left\{\left|\mathrm{D}^\alpha f(y) - \mathrm{D}^\alpha f(a)\right| : y \in [a, x],\ \alpha \text{ multi-indice tel que } |\alpha| = n\right\}$;

en particulier,

$$\lim_{x \to a} \frac{|\rho_n(x;a)|}{|x-a|^n} = 0. \tag{3.18}$$

DÉMONSTRATION

Posons

$$\varphi(t) = f(a + t\xi), \quad \xi = (x - a),\ 0 \leqslant t \leqslant 1.$$

Puisque le segment $[a, x]$ est contenu dans Ω par hypothèse, φ est bien défini dans $[0, 1]$; aussi φ est de classe C^n car $f \in C^n(\Omega)$. Notons que $\varphi(0) = f(a)$, $\varphi(1) = f(x)$; on calcule les dérivées $\varphi^{(k)}(t)$ en utilisant la formule pour la dérivation des fonctions composées ce qui donne, en écrivant $\xi = (\xi_1, \ldots, \xi_N)$, $\xi_i = x_i - a_i$,

$$\varphi'(t) = \sum_{i=1}^{N} \xi_i \,\mathrm{D}_i f(a + t\xi)$$

$$\varphi''(t) = \sum_{1 \leqslant i,\, j \leqslant N} \xi_i \xi_j \,\mathrm{D}_{ij} f(a + t\xi)$$

$$\varphi^{(k)}(t) = \sum_{1 \leqslant i_1, \ldots, i_k \leqslant N} \xi_{i_1} \xi_{i_2} \cdots \xi_{i_k} \,\mathrm{D}_{i_1 \ldots i_k} f(a + t\xi).$$

De (3.15) on voit que

$$\frac{\varphi^{(k)}(0)}{k!} = \sum_{|\alpha| = k} \frac{(x-a)^\alpha}{\alpha!} \mathrm{D}^\alpha f(a), \quad 1 \leqslant k \leqslant n$$

$$\frac{\varphi^{(n)}(\theta)}{n!} = \sum_{|\alpha| = n} \frac{(x-a)^\alpha}{\alpha!} \mathrm{D}^\alpha f\big(a + \theta(x-a)\big).$$

L'application de la formule de Taylor (3.12) pour une variable donne maintenant (3.17). De plus, avec $\alpha = (\alpha_1, \ldots, \alpha_N)$,

$$\begin{aligned} |\rho_n(x;a)| &\leqslant \sum_{|\alpha| = n} \cdot \delta(x) \cdot \frac{|x_1 - a_1|^{\alpha_1} \cdots |x_N - a_N|^{\alpha_N}}{\alpha_1! \cdots \alpha_N!} \\ &= \frac{\delta(x)}{n!}\left(|x_1 - a_1| + \cdots + |x_N - a_N|\right)^n \end{aligned}$$

en utilisant (3.16) avec $k = n$; comme

$$|x_1 - a_1| + \cdots + |x_N - a_N| \leqslant N|x - a|\,,$$

on a aussi

$$\big|\rho_n(x;a)\big| \leqslant \delta(x)|x-a|^n \cdot \frac{N^n}{n!}\,.$$

La continuité des $\mathrm{D}^\alpha f(y)$, $|\alpha| = n$, au point $y = a$ donne que $\delta(x) \to 0$ lorsque $x \to a$ d'où la relation (3.18) pour $\rho_n(x;a)$. ♦

3.3.3 Remarques

Le cas $n = 1$ de la proposition 3.3.2 était déjà vu dans la section 2.3 comme la ***formule des accroissements finis***. Soulignons ici le cas $n = 2$ de la proposition 3.3.2 qui est d'usage fréquent ; on peut l'écrire comme suit :

$$f(x) = f(a) + \nabla f(a) \cdot (x-a) + \frac{1}{2}\big\{H(x-a)\big\} \cdot (x-a) + \rho_2(x;a) \qquad (3.19)$$

où $H = \big[\mathrm{D}_{ij} f(a)\big]$, la matrice hessienne de f au point a, et $\rho_2(x;a)/|x-a|^2 \to 0$ lorsque $x \to a$, f étant supposé de classe C^2 dans une boule centrée à a.

On peut interpréter le développement limité d'ordre n en termes suivants : si f est de classe C^n dans une boule ouverte centrée à $a \in \mathbb{R}^N$, alors

$$f(x) = P_n(x;a) + \rho_n(x;a)$$

où $P_n(x;a)$ est un polynôme de degré $\leqslant n$ dans les variables $x_1, \ldots, x_N$ tel que

$$f(a) = P_n(a;a)\,, \quad \mathrm{D}^\alpha f(a) = \mathrm{D}^\alpha P_n(a;a)\,, \quad |\alpha| \leqslant n$$

et

$$\lim_{x \to a} \frac{\big|\rho_n(x;a)\big|}{|x-a|^n} = 0\,.$$

On dira que $P_n(x;a)$ est le ***polynôme de Taylor d'ordre*** n de la fonction f au point a. Il est clair que si deux fonctions f et g (de classe C^n dans une boule ouverte centrée à $a \in \mathbb{R}^N$) sont égales dans un voisinage ouvert du point a, alors leurs polynômes de Taylor d'ordre n au point a seront identiques.

3.3.4 Généralisation aux fonctions vectorielles

Une généralisation du développement limité au cas d'une fonction vectorielle $f = (f_1, \ldots, f_M)$ s'obtient immédiatement en appliquant la proposition 3.3.2 à chaque composante f_i séparément. On obtient ainsi l'énoncé suivant : *si $f \in C^n(\Omega;\mathbb{R}^M)$ et a, x sont dans Ω avec le segment $[\,a\,,\,x\,]$ contenu dans Ω alors*

$$f(x) = \big(P_{n1}(x;a), \ldots, P_{nM}(x;a)\big) + \rho_n(x;a)$$

où $P_{ni}(x;a)$ *est le polynôme de Taylor d'ordre* n *de la fonction* f_i *au point* a *et* $\rho_n = (\rho_{n1}, \ldots, \rho_{nM})$ *est tel que* $|\rho_{ni}(x;a)|/|x-a|^n = 0$ *lorsque* $x \to a$ *pour* $i = 1, \ldots, n$. On peut aussi écrire des estimations pour $|\rho_n(x;a)|$ en partant des estimations pour les $|\rho_{ni}(x;a)|$ ou en utilisant une forme intégrale pour ρ_n (Ex. 8, § 3.3.5). En écrivant

$$\mathrm{D}^\alpha f = \left(\mathrm{D}^\alpha f_1, \ldots, \mathrm{D}^\alpha f_M\right),$$

on peut même écrire

$$f(x) = f(a) + \sum_{1 \leqslant |\alpha| \leqslant n} \frac{(x-a)^\alpha}{\alpha!} \mathrm{D}^\alpha f(a) + \rho_n(x;a) \tag{3.20}$$

sous des conditions déjà énumérées : $f \in C^n(\Omega; \mathbb{R}^M)$ et Ω ouvert dans $\mathbb{R}^N$ tel que le segment $[a, x] \subset \Omega$. Par contre, la version (3.17) doit être évitée ou modifiée dans le cas vectoriel ; en effet, en appliquant (3.17) à chaque composante f_i de f on aura M nombres, non nécessairement identiques, $\theta_1, \ldots, \theta_M$, entre $]0, 1[$ et on a déjà vu, même dans le cas de $n = 1$ (théorème des accroissements finis) que ce genre de formulation est en défaut (Ex. 2, § 2.3.5).

3.3.5 Exercices

1 Soit $f(x;y) = a(x)b(y)$ où a, b sont deux fonctions réelles d'une variable réelle de classe C^n dans un intervalle ouvert contenant $x = x_0$ respectivement $y = y_0$. Vérifier que le polynôme de Taylor de f d'ordre n au point (x_0, y_0) s'obtient à partir de la multiplication des polynômes de Taylor d'ordre n de a au point x_0 et de b au point y_0 en retenant seulement les termes de types $c(x-x_0)^k(y-y_0)^\ell$, $k + \ell \leqslant n$. En déduire :

(i) le polynôme de Taylor d'ordre 4 pour $f(x,y) = \cos x \cos y$ au point $(0,0)$;
(ii) le polynôme de Taylor d'ordre 3 pour $f(x,y) = e^x \sin y$ au point $(0,0)$.

2 Démontrer que

$$\cos x \cos y = 1 - \frac{1}{2}(x^2 + y^2) + \frac{1}{24}(x^4 + 6x^2y^2 + y^4) + \rho(x,y)$$

avec

$$|\rho(x,y)| \leqslant \frac{1}{12}(|x| + |y|)^4.$$

3 Démontrer que

$$e^x \sin y = y + xy + \frac{1}{6}(3x^2y - y^3) + \rho(x,y)$$

avec

$$|\rho(x,y)| \leqslant \frac{\delta(x,y)}{6}(|x| + |y|)^3$$

$$\delta(x,y) \leqslant e^x + 1 \text{ si } x \geqslant 0, \quad \delta(x,y) \leqslant 2 \text{ si } x \leqslant 0.$$

4 Soit $f(x,y) = a(x+y)$ où a est une fonction réelle d'une variable réelle, de classe C^n, dans un intervalle ouvert contenant $t_0 = x_0 + y_0$. Ecrire le polynôme de Taylor de f d'ordre n au point (x_0, y_0) ; quel est le coefficient du terme $(x-x_0)^i(y-y_0)^{k-i}$ dans ce polynôme ?

5 Développer $f(x,y,z)$ suivant les puissances de $(x-1)$, $(y-1)$, $(z-1)$ si $f(x,y,z) = x^2 + y^2 + z^2 - 2xy - 2xz - 2yz$:

(i) directement ;
(ii) en utilisant la formule de Taylor.

6 Trouver les polynômes de Taylor d'ordre trois au point spécifié :

(i) $f(x,y) = \ln(x^2 + 2y^2) - 2x + 2$, au point $(1,0)$;
(ii) $f(x,y) = \exp(xy)$, au point $(1,1)$;
(iii) $f(x,y,z) = x\cos y + y \sin z$, au point $(0,0,0)$.

7 Soit Ω une partie ouverte de $\mathbb{R}^N$. L'application $T : C^\infty(\Omega) \to C^\infty(\Omega)$ donnée par $Tf = \sum_{\alpha \in I} a_\alpha \, \mathrm{D}^\alpha f$, où I est un ensemble fini de multi-indices et les a_α sont des nombres réels, s'appelle un ***opérateur différentiel à coefficients constants***. Vérifier que l'ensemble de ces opérateurs différentiels forme un anneau commutatif.

8 Démontrer que dans (3.20) on a

$$\rho_n(x;a) = \sum_{|\alpha|=n} \frac{(x-a)^\alpha}{\alpha!} \int_0^1 n(1-t)^{n-1} \left\{ \mathrm{D}^\alpha f\big(a + t(x-a)\big) - \mathrm{D}^\alpha f(a) \right\} \mathrm{d}t \,.$$

(Suivre la démonstration de la proposition 3.3.2 en utilisant la formule intégrale (3.12) pour le reste appliquée sur chaque composante f_i de f séparément.)

9 Démontrer (3.16) par induction sur N. (Pour $N = 1$, c'est évident ; pour $N = 2$, c'est la formule du binôme.)

3.4 MAXIMUM ET MINIMUM

On a vu dans la section 2.4 qu'une condition *nécessaire* pour qu'une fonction dérivable f possède un extremum au point a est que ce dernier soit un point critique de f, c'est-à-dire $f'(a) = 0$. Maintenant nous étudierons les points critiques de f en utilisant les dérivées d'ordre deux ; en particulier nous donnerons une condition *suffisante* sur la matrice hessienne de f pour que le point critique a de f soit un extremum local.

3.4.1 Matrices de type positif ou négatif

Soit A une matrice réelle symétrique de dimension (N, N) ; on dit que A est une matrice de ***type positif*** si $Ax \cdot x \geqslant 0$, $x \in \mathbb{R}^N$; elle s'appelle une matrice de ***type strictement positif*** si $Ax \cdot x > 0$, pour tout $x \neq 0$ dans $\mathbb{R}^N$; A s'appelle une matrice de ***type négatif*** ou ***type strictement négatif*** selon que $-A$ est une matrice de type positif ou de type strictement positif.

Rappelons que si $A = [a_{ij}]$ est une matrice (N, N) alors

$$Ax \cdot x = \sum_{1 \leqslant i,j \leqslant N} a_{ij} x_i x_j \,,$$

$Q(x) = Ax \cdot x$ s'appelle la ***forme quadratique*** associée à A. Nous utiliserons les termes « type positif », « type négatif », « forme quadratique », etc., seulement pour les matrices carrées *réelles* et *symétriques*. Si une telle matrice A est ni de type positif ni de type négatif, nous dirons que A est de ***type indéfini*** ; autrement dit, s'il existe x et y dans $\mathbb{R}^N$ tels que d'une part $Ax \cdot x > 0$ et d'autre part $Ay \cdot y < 0$, alors A sera une matrice de type indéfini. Nous dirons aussi que la forme quadratique $Q(x) = Ax \cdot x$ est « de type positif », « de type indéfini », etc., lorsque la matrice A elle-même est de « de type positif », « de type indéfini », etc.

De l'algèbre linéaire on sait qu'une matrice réelle symétrique A est diagonalisable et que toutes ses valeurs propres sont réelles ; A est de type positif (respectivement, négatif) si ses valeurs propres sont toutes $\geqslant 0$ (respectivement, $\leqslant 0$) ; elle est de type strictement positif (respectivement, strictement négatif) si ses valeurs propres sont toutes strictement positives (respectivement, strictement négatives) ; finalement, elle est de type indéfini si une de ses valeurs propres est strictement positive et qu'une autre est strictement négative.

Comme conséquence de la possibilité de la diagonalisation d'une matrice réelle symétrique A, on obtient immédiatement que la forme quadratique associée à A est une combinaison linéaire des carrés des formes linéaires :

$$Q(x) = Ax \cdot x = \sum_{i=1}^{r} c_i \varphi_i^2(x) \,, \quad x \in \mathbb{R}^N \,, \tag{3.21}$$

où les c_i sont réels, non nuls, et chaque φ_i est une forme linéaire en $x \in \mathbb{R}^N$, c'est-à-dire, il existe $v_i \in \mathbb{R}^N$ tel que

$$\varphi_i(x) = x \cdot v_i \,, \quad x \in \mathbb{R}^N \,;$$

de plus, les φ_i (c'est-à-dire les v_i) sont linéairement indépendants. En effet, soit $p_1, \ldots, p_N$ une base orthonormale de $\mathbb{R}^N$ formée des vecteurs propres de A avec

$$Ap_i = \lambda_i p_i \,, \quad i = 1, \ldots, N \,,$$

où λ_i est la valeur propre de A associée au vecteur propre p_i ; l'existence d'une telle base est équivalente à la diagonalisation de la matrice symétrique A. Alors un calcul simple montre que si

$$x = \sum_{i=1}^{N} t_i p_i\,, \quad t_i = x \cdot p_i\,,$$

on aura

$$Ax = \sum_{i=1}^{N} t_i \lambda_i p_i$$

et

$$Ax \cdot x = \sum_{i=1}^{N} \lambda_i t_i^2 = \sum_{i=1}^{N} \lambda_i \big(x \cdot p_i\big)^2\,.$$

En retenant les $\lambda_i \neq 0$ on obtient une représentation de la forme (3.21). Même si la représentation (3.21) de la forme quadratique associée à A n'est pas unique (c'est-à-dire les $c_i \neq 0$ et les φ_i linéairement indépendants peuvent être choisis de différentes manières pour une même forme quadratique $Ax \cdot x$), les quantités suivantes sont des ***invariants*** associés à A : le nombre r, le nombre n_+ des $c_i > 0$, le nombre n_- des $c_i < 0$; on peut démontrer que r est égal au ***rang*** de la matrice A ; le nombre $n_+ - n_-$ est souvent appelé la ***signature*** de A ou de la forme quadratique associée à A. Il est clair que la connaissance de r et de n_+ détermine n_- et la signature ; aussi, la connaissance de r et de la signature détermine n_+ et n_- car $n_+ + n_- = r$. On voit qu'une matrice symétrique A est de type positif si et seulement si $r = n_+$; elle est de type strictement positif si et seulement si $r = n_+ = N$ (où A est de type (N, N)). De même, A est de type négatif (ou strictement négatif) si et seulement si $r = n_-$ (ou $r = n_- = N$).

La connaissance des valeurs propres d'une matrice symétrique détermine, évidemment et immédiatement, le type de la matrice. Mais la détermination des valeurs propres d'une matrice carrée d'ordre N est assez laborieuse dès que $N \geqslant 3$; nous donnerons dans la suite un procédé simple pour la détermination du type d'une matrice symétrique.

Il est intuitivement clair de la représentation (3.21) que si A est de type strictement positif, alors il existe un $c > 0$ tel que

$$Ax \cdot x \geqslant c\|x\|^2\,, \quad x \in \mathbb{R}^N\,. \tag{3.22}$$

Comme (3.22) joue un rôle important dans les considérations qui suivent, nous formulons cela comme un lemme et en donnons une démonstration directe.

3.4.2 Lemme

Soit A une matrice carrée, réelle et symétrique, d'ordre N. Alors A est de type strictement positif si et seulement si (3.22) est vérifié pour un $c > 0$.

DÉMONSTRATION

Il est évident que (3.22) entraîne que A est de type strictement positif ; démontrons la réciproque.

Posons $Q(x) = Ax \cdot x$, $x \in \mathbb{R}^N$; il est clair que

$$Q(tx) = t^2 Q(x), \quad t \in \mathbb{R}, \ x \in \mathbb{R}^N$$

et que Q est continue dans $\mathbb{R}^N$; de plus, si A est de type strictement positif alors $Q(x) > 0$ si $x \neq 0$.

Soit

$$c = \inf_{x \in S} Q(x), \quad S = \{x \in \mathbb{R}^N : \|x\| = 1\}.$$

On voit que $c > 0$ car Q étant continue sur l'ensemble compact S, l'infimum c est atteint en un point $\theta \in S$ d'où $c = Q(\theta) > 0$.

Si $x \neq 0$, on aura, avec $t = 1/\|x\|$, que $tx \in S$ d'où

$$Ax \cdot x = Q(x) = t^{-2} Q(tx) \geqslant c\|x\|^2 ;$$

comme (3.22) est évident pour $x = 0$, la démonstration du lemme est achevée. ♦

3.4.3 Remarque

L'inégalité (3.22) reste valable pour n'importe quel choix de norme $p(x)$ en place de $\|x\|$, quitte à choisir c convenablement ; ainsi si A est de type strictement positif, on a aussi

$$Ax \cdot x \geqslant c|x|^2, \quad x \in \mathbb{R}^N$$

pour un $c > 0$ approprié. En fait, on peut choisir le même c que dans (3.22) car $\|x\| \geqslant |x|$. Le cas général d'une norme $p(x)$ quelconque s'obtient directement de notre démonstration ou à partir de l'exercice 5 (§ 1.7.4) concernant deux normes quelconques dans $\mathbb{R}^N$.

3.4.4 Proposition

Soit $f \in C^2(\Omega)$, Ω une partie ouverte de $\mathbb{R}^N$.

(a) Si $a \in \Omega$ est un point de minimum relatif (respectivement, maximum relatif) pour f alors $f'(a) = 0$ et H, la matrice hessienne de f au point a, est de type positif (respectivement, de type négatif).

(b) Soit $a \in \Omega$ tel que $f'(a) = 0$; soit H la matrice hessienne de f au point a.

(i) Si H est de type strictement positif alors a est un point de minimum relatif strict pour f.

(ii) Si H est de type strictement négatif alors a est un point de maximum relatif strict pour f.

(iii) Si H est de type indéfini alors a n'est pas un point d'extremum relatif pour f.

DÉMONSTRATION

(a) Supposons que a est un point du minimum local pour f ; alors nous savons que $f'(a) = 0$ (prop. 2.4.1) ; du développement limité d'ordre 2 (prop. 3.3.2) de f au point a on a, compte tenu que $f'(a) = 0$, que (cf. (3.19)) :

$$f(a+x) = f(a) + \frac{1}{2} Hx \cdot x + \rho(x) \tag{3.23}$$

où

$$\lim_{x \to 0} \frac{\rho(x)}{|x|^2} = 0 .$$

Puisque a est un point du minimum relatif pour f, il existe un $r > 0$ tel que

$$f(a+x) \geqslant f(a) \quad \text{si } |x| < r$$

d'où, grâce à (3.23), on aura

$$\frac{1}{2} Hx \cdot x + \rho(x) \geqslant 0 \quad \text{si } |x| < r .$$

En passant de x à tx (t réel, suffisamment petit, x fixe quelconque) dans cette dernière inégalité, on aura

$$\frac{t^2}{2} Hx \cdot x + \rho(tx) \geqslant 0 ,$$

d'où

$$Hx \cdot x \geqslant \lim_{t \to 0} \frac{-2\rho(tx)}{t^2} = 0$$

à cause de la propriété de ρ mentionnée sous (3.23) ce qui démontre que H est de type positif.

Si a est un point du maximum relatif pour f alors c'est un point du minimum relatif pour $-f$; comme la hessienne de $-f$ au point a est $-H$, on déduit de ce qui précède que $-H$ est de type positif d'où H est de type négatif.

(b) Si $f'(a) = 0$, on a, de nouveau, (3.23) avec

$$\rho(x) = |x|^2 \varepsilon(x) , \quad \varepsilon(x) \to 0 \text{ lorsque } x \to 0 . \tag{3.24}$$

(i) Si H est de type strictement positif, fixons un $c > 0$ tel que

$$Hx \cdot x > c|x|^2 , \quad x \in \mathbb{R}^N .$$

On aura alors de (3.23)

$$f(a+x) - f(a) \geqslant \left\{\frac{c}{2} - \varepsilon(x)\right\} |x|^2 , \quad x \in \mathbb{R}^N ;$$

puisque $\varepsilon(x) \to 0$ lorsque $x \to 0$, il existe un $r > 0$ tel que

$$|\varepsilon(x)| < \frac{c}{2} \quad \text{si } |x| < r$$

d'où

$$f(a+x) - f(a) > 0 \quad \text{si } |x| < r ;$$

ce qui montre que a est un point du minimum relatif strict pour f.

(ii) La démonstration ici s'obtient de (i) par un passage à $-f$.

(iii) Si H est de type indéfini, nous pouvons trouver deux éléments ξ, η dans $\mathbb{R}^N$ tels que

$$H\xi \cdot \xi = \alpha > 0 , \quad H\eta \cdot \eta = \beta < 0 , \quad |\xi| = |\eta| = 1 .$$

De (3.23) et (3.24) on aura

$$f(a+t\xi) - f(a) = t^2 \left\{\frac{\alpha}{2} + \varepsilon(t\xi)\right\} > 0$$

si t est réel et suffisamment petit car $\varepsilon(t\xi) \to 0$ lorsque $t \to 0$; de même,

$$f(a+t\eta) - f(a) = t^2 \left\{\frac{\beta}{2} + \varepsilon(t\xi)\right\} < 0$$

si t est réel et suffisamment petit. Ainsi, dans tout voisinage du point a, il y a des points (de type $a + t\xi$) où les valeurs de f seront plus grandes que $f(a)$ et il y a aussi des points (de type $a + t\eta$) où les valeurs de f sont plus petites que $f(a)$; on conclut que a n'est pas un point d'extremum relatif pour f. ♦

3.4.5 Remarque

Un point a du type envisagé en (b)(iii) de la proposition 3.4.4 s'appelle souvent un ***point de selle*** (ou ***point-selle***) de f. Plus précisément, un point critique a est un point-selle de f si dans tout voisinage ouvert V de a il existe des points x, y tels que d'une part $f(x) > f(a)$ et d'autre part $f(y) < f(a)$. On peut donc reformuler (b)(iii) de la proposition 3.4.4 en disant que si $f \in C^2(\Omega)$, $f'(a) = 0$, alors a est un point-selle de f si la hessienne de f au point a est de type indéfini.

3.4.6 Une méthode pour la vérification de la positivité d'une matrice

Soit A une matrice carrée, réelle et symétrique. La méthode mentionnée en titre de ce paragraphe consiste à faire des « complétion des carrées » dans la forme quadratique associée à A. Nous illustrons la méthode générale par des exemples.

Soit

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 6 \end{bmatrix}$$

La forme quadratique associée à A est, avec $x = (x_1, x_2, x_3)$,

$$Q(x) = Ax \cdot x = x_1^2 + 3x_2^2 + 6x_3^2 + 2x_1x_2 - 2x_1x_3 + 2x_2x_3 \,.$$

Dans une première étape, on écrit

$$Q(x) = \big(x_1 + x_2 - x_3\big)^2 + 2x_2^2 + 5x_3^2 + 4x_2x_3 \,.$$

L'avantage de cette écriture est que $Q(x)$ est maintenant la somme d'un carré d'une forme linéaire et d'une forme quadratique avec moins de variables (ici, deux : x_2, x_3). En continuant ainsi, on aura

$$Q(x) = \big(x_1 + x_2 - x_3\big)^2 + 2\big(x_2 + x_3\big)^2 + 3x_3^2$$

une représentation du type (3.21) ; cette dernière montre clairement que $Q(x) \geqslant 0$ pour $x \in \mathbb{R}^3$; de plus, si $Q(x) = 0$, alors

$$x_1 + x_2 - x_3 = 0\,, \quad x_2 + x_3 = 0\,, \quad x_3 = 0$$

ce qui n'est possible que si $x = 0$; donc $Q(x) > 0$ si $x \neq 0$ et la matrice A est de type strictement positif. Notons que les valeurs propres de A seront difficiles à calculer exactement ; elles sont, approximativement,

$$0,2823\,; \quad 3,3254\,; \quad 6,3923$$

solutions de l'équation polynomiale

$$\det(A - \lambda I) = -\lambda^3 + 10\,\lambda^2 - 24\,\lambda + 6 = 0\,.$$

En fait, on voit immédiatement sur l'équation polynomiale que $\lambda \leqslant 0$ ne peut pas être une racine.

Le procédé présenté ci-dessus, s'il est complètement applicable, donnera successivement les formes linéaires φ_1, φ_2, ... chacune contenant au moins une variable de moins avec

$$Q(x) = \sum_{i=1}^{r} c_i \varphi_i^2(x)\,, \quad x \in \mathbb{R}^N \;;$$

l'indépendance linéaire des φ_i est parfaitement évidente.

Il se peut qu'à un point des calculs, les carrés des variables n'apparaissent plus. Illustrons ceci par un exemple simple :

$$x = (x_1, x_2)\,, \quad A = \begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix}\,, \quad Q(x) = Ax \cdot x = x_1x_2 \,.$$

Ici, on peut écrire

$$Q(x) = \frac{1}{4}\{(x_1 + x_2)^2 - (x_1 - x_2)^2\}$$

d'où on voit que A est de type indéfini :

$$Q(1,1) > 0\,, \quad Q(1,-1) < 0\,.$$

Illustrons ce procédé par un exemple général. Si $Ax \cdot x = Q(x)$, $x = (x_1, \ldots, x_N)$, ne contient aucun terme de type cx_i^2, alors la matrice $A = [a_{ij}]$ est telle que $a_{ii} = 0$ pour tout i ; si $a_{ij} \neq 0$, $i \neq j$, on aura, on posant $a = a_{ij}$,

$$Q(x) = ax_ix_j + x_i\varphi(x) + x_j\psi(x) + R(x)$$

où φ, ψ sont des formes linéaires en x, ne contenant pas x_i, x_j, et R est une forme quadratique indépendante de x_i, x_j. Alors,

$$\begin{aligned} Q(x) &= a\left\{x_i + \frac{\psi(x)}{a}\right\}\left\{x_j + \frac{\varphi(x)}{a}\right\} - \frac{\varphi(x)\psi(x)}{a} + R(x) \\ &= \frac{a}{4}\{\varphi_1^2(x) + \varphi_2^2(x)\} + Q_1(x) \end{aligned}$$

où

$$\begin{aligned} \varphi_1(x) &= x_i + x_j + \frac{1}{2}\{\varphi(x) + \psi(x)\} \\ \varphi_2(x) &= x_i - x_j + \frac{1}{2}\{\psi(x) - \varphi(x)\} \\ Q_1(x) &= R(x) - \frac{\varphi(x)\psi(x)}{a}\,. \end{aligned}$$

Notons que les formes linéaires φ_1, φ_2 sont linéairement indépendantes et Q_1 est une forme quadratique ne contenant pas les variables x_i, x_j. On peut maintenant continuer le procédé avec Q_1 mais d'ores et déjà il est clair que la forme quadratique Q est de type indéfini. On voit donc que le procédé de « complétion des carrés » permet de décider assez rapidement si oui ou non une matrice est de type positif, négatif, strictement positif, strictement négatif ou indéfini.

Un critère de positivité ou de négativité stricte qui est utile est le suivant : *soit $A = [a_{ij}]$ une matrice carrée, réelle et symétrique de dimension N ; soient $d_1, d_2, \ldots, d_N$ les N* ***mineurs principaux diagonaux****, c'est-à-dire*

$$d_1 = a_{11}\,,\ d_2 = \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\,,\ \ldots\,,\ d_N = \det A\,;$$

alors A est de type strictement positif si et seulement si

$$d_1 > 0\,,\ d_2 > 0\,,\ \ldots\,,\ d_N > 0$$

et A est de type strictement négatif si et seulement si $-A$ est strictement positif ce qui donne

$$d_1 < 0\,,\ d_2 > 0\,,\ d_3 < 0\,,\ \ldots\,,\ (-1)^N d_N > 0\,.$$

Noter que si les nombres $d_i \geqslant 0$, on ne peut pas conclure, en général, que la matrice est de type positif (Ex. 3, § 3.4.8). Par contre, on peut démontrer le théorème suivant : *soit* $d(i_1, \ldots, i_p)$ *le déterminant de la matrice obtenue à partir de* A *en retenant seulement les lignes et colonnes numéros* $i_1, \ldots, i_p$, $1 \leqslant i_1 < \ldots < i_p \leqslant N$; *alors* A *est de type positif si et seulement si, pour tout* $1 \leqslant i_1 < \ldots < i_p \leqslant N$,

$$d(i_1, \ldots, i_p) \geqslant 0\,, \quad p = 1, 2, \ldots, N$$

et A *est de type négatif si et seulement si*

$$(-1)^p d(i_1, \ldots, i_p) \geqslant 0\,, \quad p = 1, 2, \ldots, N$$

pour tout $1 \leqslant i_1 < \ldots < i_p \leqslant N$, *ce qui résulte du fait que* A *est de type négatif si et seulement si* $-A$ *est de type positif*.

Remarquer que dans ce dernier critère on doit calculer $2^N - 1$ déterminants, alors que dans le critère précédent on en calcul seulement N. Les quantités $d(i_1, \ldots, i_p)$ s'appellent aussi les ***mineurs principaux*** ; avec cette notation,

$$d_1 = d(1)\,,\ d_2 = d(1,2)\,,\ d_3 = d(1,2,3)\,,\ \ldots$$

Pour une discussion détaillée de ces critères voir Gantmacher [21 (vol. 1), pp. 306-307].

3.4.7 Cas particulier de deux variables

Ici nous spécialisons nos discussions précédentes au cas d'une fonction réelle f de deux variables réelles notées x, y ; on suppose que f est de classe C^2 dans une partie ouverte fixe de $\mathbb{R}^2$. De telles fonctions sont importantes, entre autre dans l'étude des surfaces dans $\mathbb{R}^3$. Nous interpréterons les résultats analytiques dans la suite en termes géométriques.

Supposons que (a, b) est un point critique de f, c'est-à-dire

$$\mathrm{D}_1 f(a,b) = 0\,, \quad \mathrm{D}_2 f(a,b) = 0\,.$$

Pour étudier la nature de ce point critique, nous calculons la matrice hessienne H de f au point (a, b) ; supposons que

$$H = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}, \quad \begin{aligned} \alpha &= \mathrm{D}_{11} f(a,b)\,, \quad \gamma = \mathrm{D}_{22} f(a,b) \\ \beta &= \mathrm{D}_{12} f(a,b) = \mathrm{D}_{21} f(a,b)\,. \end{aligned}$$

Alors, la forme quadratique associée à H est

$$Q(x,y) = \alpha x^2 + 2\beta xy + \gamma y^2$$

Si $\alpha \neq 0$, on aura

$$Q(x,y) = \alpha \left(x + \frac{y\beta}{\alpha} \right)^2 + (\alpha\gamma - \beta^2)\frac{y^2}{\alpha}\,.$$

Posons $D = \alpha\gamma - \beta^2 = \det(H)$; D s'appelle le ***discriminant*** de la forme quadratique Q. On voit que :

- si $\alpha > 0$, $D > 0$ alors Q est de type strictement positif ;
- si $\alpha < 0$, $D > 0$ alors Q est de type strictement négatif ;
- si $D < 0$, $\alpha \neq 0$ alors Q est de type indéfini (car les signes de α et D/α sont différents).

Notons que $D > 0$ entraîne que α, γ sont non nuls. De l'analyse précédente on conclut comme suit :

(i) si $D > 0$, le point critique est un minimum relatif strict si $\alpha > 0$, un maximum relatif strict si $\alpha < 0$;

(ii) si $D < 0$, le point critique est un point-selle.

Ce dernier point était établi sous l'hypothèse $\alpha \neq 0$; si $\alpha = 0$ mais $\gamma \neq 0$, $D < 0$ (donc $\beta \neq 0$), on peut raisonner sur

$$Q(x,y) = \gamma\left(y + \frac{\beta x}{\gamma}\right)^2 + \frac{(\alpha\gamma - \beta^2)x^2}{\gamma}$$

ce qui montre que dans ce cas Q est de type indéfini (car les signes de γ et D/γ sont différents) ; finalement, si $\alpha = \gamma = 0$ mais $D < 0$ (donc $\beta \neq 0$), on aura

$$Q(x,y) = 2\beta xy = \frac{\beta}{2}\left\{(x+y)^2 - (x-y)^2\right\}$$

ce qui est, de nouveau, une forme quadratique de type indéfini ; ceci établit (ii) complètement.

Il reste à voir le cas $D = 0$; dans ce cas on vérifie, à partir des calculs précédents, les affirmations suivantes :

(iii) si $D = 0$ alors

Q est de type positif (non strictement) si $\alpha > 0$ ou si $\alpha = 0$, $\gamma \geqslant 0$;

Q est de type négatif (non strictement) si $\alpha < 0$ ou si $\alpha = 0$, $\gamma \leqslant 0$.

Mais ces affirmations dans le cas $D = 0$ ne permettent pas de déterminer la nature exacte du point critique (a,b) seulement sur la base de la matrice hessienne H et la proposition 3.4.4 n'est pas applicable. Illustrons ceci par deux exemples simples.

Soit $f(x,y) = x^2$; au point critique $(0,0)$ la matrice hessienne H de f est

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

dont le déterminant $D = 0$; ici $(0,0)$ est évidemment un point de minimum absolu (ainsi que minimum relatif non strict). Par contre, si $f(x,y) = x^3$, la matrice hessienne de f au point critique $(0,0)$ est la matrice nulle ; ici $(0,0)$ est un point-selle. D'autre exemples sont donnés dans les exercices (§ 3.4.8).

Si le déterminant de la matrice hessienne de la fonction (de N variables) est nul à un point critique, ce dernier s'appelle un point critique ***dégénéré***, autrement il s'appelle un point critique ***non dégénéré***. Dans le cas des points critiques non dégénérés des fonctions de deux variables, les affirmations (i) et (ii) précédentes décident complètement l'extrémalité ou non du point critique. Par contre, un point critique dégénéré, même dans le cas d'une fonction de deux variables, exige une étude plus détaillée pour déterminer son extrémalité.

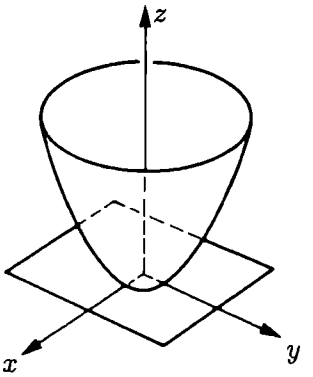
Une interprétation géométrique de la nature d'un point critique est facile à donner. Considérons la surface S d'équation cartésienne $z = f(x, y)$ dans $\mathbb{R}^3$. Soit (a, b) un point critique de f ; alors le plan tangent à la surface S au point $P = (a, b, c)$, $c = f(a, b)$, est un plan parallèle au plan x, y ; son équation cartésienne sera $z = c$. Si le point critique (a, b) est un point de minimum relatif pour f, alors le point P est un *creux* local de S ; s'il est un point de maximum relatif pour f, le point P est une *bosse* locale de S ; si le point critique est un point de selle pour f, alors la surface S aura un *col* à ce point.

Les exemples typiques de ces trois cas sont fournis par les fonctions suivantes :

(1) $f(x, y) = x^2 + y^2$;
(2) $f(x, y) = -(x^2 + y^2)$;
(3) $f(x, y) = x^2 - y^2$.

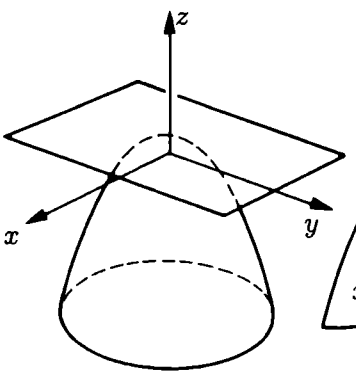
Le point $(0, 0)$ est un point critique pour toutes les trois fonctions ; il est un point de minimum relatif pour (1), un point de maximum relatif pour (2) et un point de selle pour (3). Géométriquement, ces fonctions représentent respectivement

(1) un paraboloïde de révolution ayant l'axe des z pour axe de symétrie et tourné vers « le haut » ;
(2) le même paraboloïde tourné vers « le bas » ;
(3) un paraboloïde hyperbolique équilatère.



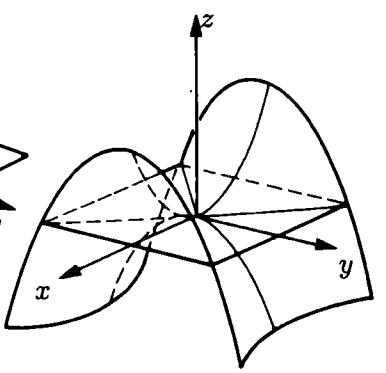
$f(x, y) = x^2 + y^2$

Fig 3.1



$f(x, y) = -x^2 - y^2$

Fig 3.2



$f(x, y) = x^2 - y^2$

Fig 3.3

3.4.8 Exercices

[1] Déterminer le type des formes quadratiques suivantes en précisant chaque fois les matrices symétriques associées :

(i) $x_1^2 + 2x_2^2 + 6x_3^2 - 2x_1x_2 - 6x_2x_3 + 4x_1x_3$;
(ii) $4x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 - 4x_2x_3 - 4x_1x_3$;
(iii) $4x_1^2 + 9x_2^2 + 2x_3^2 - 12\,x_1x_2 - 6x_2x_3 + 4x_1x_3$;
(iv) $x_2^2 + x_1x_3$;
(v) $4x_1^2 - x_2^2 - x_4^2 + 3x_2x_3 - 12\,x_1x_4$.

[2] Soit A une matrice symétrique non nulle de dimension N telle que $\operatorname{Tr} A = 0$; démontrer que A est de type indéfini.

[3] Soit

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Vérifier que les mineurs principaux diagonaux de A sont $d_1 = 0$, $d_2 = 0$, $d_3 = 1$ donc tous $\geqslant 0$, mais que A est une matrice de type indéfini.

[4] Déterminer les points critiques des fonctions définies par les expressions suivantes et étudier leur extrémalité :

(i) $x^2 + ax^2y + y^4$ $(a \in \mathbb{R})$;
(ii) $x^3 + 2y^3 - xy$;
(iii) $x^3 + y^3 - 3xy$;
(iv) $x^4 + y^4 - 4xy$;
(v) $x^n + y^n - nxy$, n entier $\geqslant 1$;
(vi) $ax^2 + 2bxy + cy^2$ $(ac \neq 0,\ a,\ c$ réels) ;
(vii) $xy(1 - x - y)$;
(viii) $(x + y)/(x^2 + y^2 + 1)$.

3.5 COORDONNÉES POLAIRES DANS $\mathbb{R}^2$

3.5.1 Introduction

Soit $(x, y) \in \mathbb{R}^2$; en partant de la trigonométrie élémentaire, nous pouvons associer à (x, y) un nombre $r \geqslant 0$ et un nombre θ tels que

$$x = r\cos\theta\,, \quad y = r\sin\theta\,. \tag{3.25}$$

Il est clair que

$$r = \left(x^2 + y^2\right)^{1/2}. \tag{3.26}$$

La valeur de θ n'est pas uniquement déterminée par (x,y) ; si $(x,y) = (0,0)$, θ reste complètement indéterminé ; si $(x,y) \neq (0,0)$, le nombre θ est déterminé à une addition de $2n\pi$ près, $n \in \mathbb{Z}$. Les couples (r,θ), $r > 0$, associés au point $(x,y) \neq (0,0)$ selon (3.25) et (3.26), s'appellent les ***coordonnées polaires*** de (x,y) ; cette terminologie sera utilisée seulement pour $(x,y) \neq (0,0)$. Le nombre r associé à (x,y) s'appelle la ***distance polaire*** de (x,y) ; le nombre θ mesure l'***angle polaire*** de (x,y).

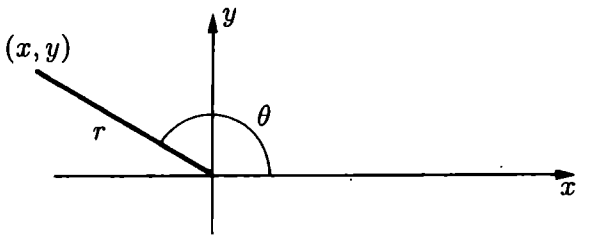


Fig. 3.4

En exigeant, arbitrairement, que la valeur de θ soit dans l'intervalle $]-\pi, \pi]$, on détermine le nombre θ uniquement, pour chaque $(x,y) \neq (0,0)$, via les formules (3.25) et (3.26) ; nous noterons cette valeur de θ par $\arg(x,y)$, l'***argument principal*** de (x,y) où la ***détermination principale de l'angle polaire*** associé à (x,y). Ainsi, avec $\theta(x,y) = \arg(x,y)$, on aura

$$\theta(x,0) = 0 \text{ si } x > 0\,; \qquad \theta(x,0) = \pi \text{ si } x < 0$$
$$\theta(0,y) = \frac{\pi}{2} \text{ si } y > 0\,; \qquad \theta(0,y) = -\frac{\pi}{2} \text{ si } y < 0$$
$$\theta(x,x) = \frac{\pi}{4} \text{ si } x > 0\,; \qquad \theta(x,x) = -\frac{3\pi}{4} \text{ si } x < 0\,.$$

Si (x,y) n'est pas de la forme $(x,0)$, $x \leqslant 0$, $\theta = \arg(x,y)$ sera strictement entre $-\pi$ et π ; alors, on aura de (3.25)

$$\operatorname{tg}\frac{\theta}{2} = \frac{\sin\theta/2}{\cos\theta/2} = \frac{2\sin\theta/2\cos\theta/2}{2\cos^2\theta/2} = \frac{\sin\theta}{1+\cos\theta} = \frac{y}{x+r}$$

d'où

$$\theta = 2\operatorname{arc\,tg}\frac{y}{x+r} \tag{3.27}$$

r étant donné par (3.26).

Les propriétés analytiques des fonctions r et arg sont résumées dans la proposition suivante.

3.5.2 Proposition

(i) La fonction r définie par (3.25) est continue partout et de classe C^∞ dans $\mathbb{R}^2 \setminus \{(0,0)\}$; la fonction $\ln r$ est harmonique dans $\mathbb{R}^2 \setminus \{(0,0)\}$.

(ii) La fonction $\theta = \arg$ définie par (3.27) est de classe C^∞ dans $\mathbb{R}^c_-$, $\mathbb{R}_- = \{(x,0) : x \leqslant 0\}$; elle est harmonique dans $\mathbb{R}^c_-$; elle ne peut pas être prolongée par continuité à un point quelconque de $\mathbb{R}^c_-$.

DÉMONSTRATION

(i) Résulte d'une vérification directe que nous laissons comme un exercice.
(ii) Notons que, dans $\mathbb{R}^c_-$,

$$\frac{\partial\theta}{\partial x} = -yr^{-2} = \frac{-y}{x^2+y^2}, \quad \frac{\partial\theta}{\partial y} = xr^{-2} = \frac{x}{x^2+y^2} \tag{3.28}$$

d'où l'on obtient que

$$\triangle\theta = \frac{\partial^2\theta}{\partial x^2} + \frac{\partial^2\theta}{\partial y^2} = 0 \quad \text{dans } \mathbb{R}^c_- .$$

Aussi, pour $x < 0$, on voit de (3.27) que

$$\lim_{\substack{y\to 0\\ y>0}} \theta(x,y) = \pi, \quad \lim_{\substack{y\to 0\\ y<0}} \theta(x,y) = -\pi$$

(relations évidentes géométriquement) ce qui montre l'impossibilité du prolongement de θ par continuité. ♦

3.5.3 Remarques

La formule suivante pour $\theta = \arg(x,y)$ est aussi utile ; elle part de l'observation que $\operatorname{tg}\theta = y/x$ si $x \neq 0$ ce qui donne

$$\theta = \operatorname{arc\,tg}\left(\frac{y}{x}\right) \tag{3.29}$$

valable *seulement* pour (x,y) tel que $x > 0$. La formule complète pour $\theta(x,y) = \arg(x,y)$ dans cette forme est la suivante : si $(x,y) \in \mathbb{R}^c_-$

$$\theta(x,y) = \begin{cases} \operatorname{arc\,tg}\dfrac{y}{x} & \text{si } x > 0 \\ \operatorname{arc\,tg}\dfrac{y}{x} + \pi & \text{si } x < 0,\ y > 0 \\ \operatorname{arc\,tg}\dfrac{y}{x} - \pi & \text{si } x < 0,\ y < 0 \\ \dfrac{\pi}{2} & \text{si } x = 0,\ y > 0 \\ -\dfrac{\pi}{2} & \text{si } x = 0,\ y < 0. \end{cases} \tag{3.30}$$

Il n'est pas a priori évident que (3.27) et (3.30) définissent la même fonction ; c'est néanmoins vrai comme indiqué dans l'exercice 2 (§ 3.5.4).

On a défini $\arg(x,y)$ en choisissant la valeur de θ, donnée par (3.25), dans l'intervalle $]-\pi, \pi]$. On aurait pu choisir la valeur de θ dans un intervalle $]a, a+2\pi]$ ou $[a, a+2\pi[$ de longueur 2π ; alors θ aurait des discontinuités le long d'une demi-droite $d_p = \{tp : t \geqslant 0\}$ où $p \in C = \{(x,y) : x^2+y^2 = 1\}$,

le cercle unité, tel que $\arg p - a$ est un multiple de 2π. Si $p = (-1,0)$, d_p devient $\mathbb{R}_-$ et on obtient la fonction arg ayant des valeurs dans $]-\pi\,,\,\pi]$. Le cas $p = (1,0)$ donne un choix de θ dans $[0\,,\,2\pi[$ (Ex. 4, § 3.5.4). Que l'intervalle choisi soit fermé à gauche ou à droite n'a aucune importance ; la continuité de la fonction dans $\mathbb{R}^2 \setminus d_p$ fixe ces valeurs dans l'intervalle ouvert correspondant.

Est-ce que l'on peut définir $\theta = \theta(x,y)$ (satisfaisant (3.25)) continûment pour tout (x,y) dans $\mathbb{R}^2 \setminus \{(0,0)\}$? La réponse est négative. En effet, une telle fonction θ donnerait une application continue et injective de C, le cercle unité, dans $\mathbb{R}$ ce qui est impossible (exercice).

Par contre, si $p \in C$, $C \setminus \{p\}$ est homéomorphe à $\mathbb{R}$, donc homéomorphe à un intervalle ouvert quelconque de $\mathbb{R}$ (Ex. 12, sect. 1.10). Si $p = (-1,0)$, l'application

$$t \longmapsto (\cos t, \sin t)\,, \quad -\pi < t < \pi$$

donne un homéomorphisme explicite simple entre $]-\pi\,,\,\pi[$ et $C \setminus \{p\}$. C'est cet homéomorphisme qui est à la base de la définition de la fonction arg.

Finalement, la fonction

$$f(r,\theta) = (r\cos\theta\,,\,r\sin\theta) \tag{3.31}$$

donne un homéomorphisme comme suit :

$$f :]0\,,\,\infty[\times]-\pi\,,\,\pi[\longrightarrow \mathbb{R}^c_- .$$

En effet, f est continue et bijective, la fonction $f^{-1} = g$ étant définie explicitement comme

$$g(x,y) = \big(r(x,y)\,,\,\arg(x,y)\big)$$

ce qui montre que $f^{-1} = g$ est continue.

Certains auteurs permettent les valeurs négatives à r dans (3.25) en prenant $\theta + \pi$ au lieu de θ ; par exemple, on a

$$\big(r\cos\,,\,r\sin\theta\big) = \big(-r\cos(\theta+\pi)\,,\,-r\sin(\theta+\pi)\big)\,.$$

Selon cet usage, (r,θ) et $(-r,\theta+\pi)$ sont les coordonnées polaires d'un même point. Malgré quelques avantages de cette façon de faire, nous ne l'utiliserons jamais dans ce livre ; pour nous $r = r(x,y)$ sera toujours $\geqslant 0$.

Une étude purement analytique de la fonction arg, évitant toute référence à la trigonométrie élémentaire, sera entreprise dans la partie du livre consacrée à l'analyse complexe. On verra que $\arg(x,y)$ est la partie imaginaire de $\ln z$, $z = x + iy$, ce qui donnera une autre explication pour le fait que arg est une fonction harmonique.

3.5.4 Exercices

1 Vérifier en détail que $\theta = \arg$ définie par (3.27) est une fonction harmonique dans $\mathbb{R}^c_-$.

2 Démontrer que (3.27) et (3.30) définissant la même fonction dans $\mathbb{R}^c_-$. (Vérifier que les deux fonctions ont le même gradient dans le domaine $\mathbb{R}^c_-$ et qu'elles coïncident pour $(x,0)$, $x > 0$.)

3 Soit

$$\Omega = \{(x,y) : c_1 < r(x,y) < c_2\} \setminus \mathbb{R}_-$$

où $0 < c_1 < c_2 < \infty$; posons $\theta(x,y) = \arg(x,y)$ comme défini par (3.27). Vérifier que $\operatorname{grad}\theta$ est borné dans Ω mais que θ n'est pas lipschitzienne dans Ω. (Cet exemple était signalé dans § 2.3.4 ; si θ était lipschitzienne dans Ω, θ serait prolongeable par continuité aux points de $\mathbb{R}_-$ dans Ω ce qui est exclu.)

4 Pour $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, posons

$$\theta(x,y) = \begin{cases} \arg(x,y) & \text{si } y \geqslant 0 \\ \arg(x,y) + 2\pi & \text{si } y < 0. \end{cases}$$

Démontrer que θ est de classe C^∞ et harmonique dans $\mathbb{R}^c_+$ ou $\mathbb{R}_+ = \{(x,0) : x \geqslant 0\}$ mais θ est discontinue à chaque point de $\mathbb{R}^*_+$. Avec cette fonction θ, vérifier que f, définie par (3.31), donne un homéomorphisme entre $]\,0\,,\infty[\,\times\,]\,0\,,2\pi\,[$ et $\mathbb{R}^c_+$.

5 Voir que f, définie par (3.31), est dérivable et même de classe C^∞ dans $]\,0\,,\infty[\,\times\mathbb{R}$ et calculer la matrice $f'(r,\theta)$; vérifier que $\det f'(r,\theta) = r$.

6 Voir que si $\theta(x,y) = \arg(x,y)$, θ est une fonction homogène de degré 0 ; vérifier directement que

$$x\frac{\partial\theta}{\partial x} + y\frac{\partial\theta}{\partial y} = 0 \quad \text{dans } \mathbb{R}^c_- ,$$

un cas particulier du théorème d'Euler (Ex. 9, § 2.2.4).

3.6 FONCTIONS HARMONIQUES

La proposition suivante donne une propriété importante des fonctions harmoniques.

3.6.1 Proposition (*principe du maximum faible*)

Soit Ω une partie ouverte bornée non vide dans $\mathbb{R}^N$; soit $f : \overline{\Omega} \to \mathbb{R}$ une fonction continue qui est harmonique dans Ω. Si $f(x) \leqslant M$ pour $x \in \operatorname{Fr}\Omega$ (M une constante réelle) alors $f(x) \leqslant M$ pour $x \in \overline{\Omega}$.

DÉMONSTRATION

Soit $\varepsilon > 0$ quelconque et posons

$$F(x) = f(x) + \varepsilon x_1^2, \quad x = (x_1, \ldots, x_N).$$

Puisque $\overline{\Omega}$ est compact et F est continue dans $\overline{\Omega}$, il existe un point $p \in \overline{\Omega}$ tel que

$$F(p) = \sup_{p \in \overline{\Omega}} F(p).$$

Démontrons que $p \in \operatorname{Fr} \Omega$; sinon, $p = (p_1, \ldots, p_N)$ serait un point de Ω et un point de maximum relatif pour F d'où, en considérant les N fonctions d'une variable

$$x_1 \mapsto F(x_1, p_2, \ldots, p_N), \ \ldots, \ x_N \mapsto F(p_1, p_2, \ldots, x_N),$$

on aurait

$$\mathrm{D}_i F(p) = 0, \quad \mathrm{D}_{ii} F(p) \leqslant 0, \quad i = 1, \ldots, N.$$

On en déduirait que $\Delta F(p) \leqslant 0$, ce qui contredirait

$$\Delta F(p) = \Delta f(p) + 2\varepsilon = 2\varepsilon > 0.$$

Si $C = \max_{x \in \overline{\Omega}} |x_1|$, on a

$$F(x) \leqslant F(p) = f(p) + \varepsilon p_1^2 \leqslant M + \varepsilon C^2, \quad x \in \overline{\Omega},$$

car $p \in \operatorname{Fr} \Omega$ et $f(p) \leqslant M$; on en déduit que

$$f(x) \leqslant M + \varepsilon(C^2 - x_1^2), \quad x \in \overline{\Omega},$$

d'où, en laissant $\varepsilon \to 0$, on aura le résultat. ♦

3.6.2 Corollaire (*principe du minimum faible*)

Avec la notation de la proposition 3.6.1, si $f(x) \geqslant m$ pour $x \in \operatorname{Fr} \Omega$ (m une constante réelle) alors $f(x) \geqslant m$ pour $x \in \overline{\Omega}$. Aussi,

$$\sup_{x \in \overline{\Omega}} |f(x)| = \sup_{x \in \operatorname{Fr} \Omega} |f(x)|.$$

La démonstration est immédiate en appliquant la proposition 3.6.1 à la fonction $-f$.

3.6.3 Remarques

Le ***principe du maximum*** que nous démontrerons plus tard affirme que si Ω est un domaine borné dans $\mathbb{R}^N$ et $f = \overline{\Omega} \to \mathbb{R}$ est continue et non constante, harmonique dans Ω, alors le supremum de $f(x)$, $x \in \overline{\Omega}$, est atteint

seulement aux points $x \in \operatorname{Fr}\Omega$. En passant à $-f$ on peut obtenir le ***principe du minimum*** qui dit que l'infimum de f n'est atteint qu'à un point de $\operatorname{Fr}\Omega$ (f étant harmonique dans Ω et continue, non constante dans $\overline{\Omega}$).

Une interprétation physique du principe du maximum ou minimum est instructive. Considérons $\overline{\Omega}$ comme un corps homogène, conducteur de chaleur, dans l'espace physique modélisé par $\mathbb{R}^3$. Une fonction continue $f : \overline{\Omega} \to \mathbb{R}$, harmonique dans Ω, avec $f = \varphi$ sur $\operatorname{Fr}\Omega$, peut être considérée comme la distribution de température dans le corps $\overline{\Omega}$ lorsque la température sur la surface $\operatorname{Fr}\Omega$ est maintenue à $\varphi(x)$ au point $x \in \operatorname{Fr}\Omega$. On suppose que $\overline{\Omega}$ ne contient aucune source où la chaleur est engendrée ou perdue et que $\overline{\Omega}$ est dans un état d'équilibre thermique. Il est alors physiquement évident qu'à cause de l'équilibre thermique, f ne peut avoir un point de maximum ou minimum en un point intérieur de Ω.

Notons que la proposition 3.6.1 reste valable sous la condition plus générale que $\Delta f(x) \geqslant 0$, $x \in \Omega$, f étant de classe C^2 dans Ω ; une telle fonction s'appelle une fonction ***sous-harmonique***. De même le corollaire 3.6.2 est valable pour une fonction ***sur-harmonique*** f qui est de classe C^2 dans Ω avec $\Delta f(x) \leqslant 0$, $x \in \Omega$. Il est clair que si f est sur-harmonique, $-f$ sera sous-harmonique et que f est harmonique si et seulement si f est à la fois sous-harmonique et sur-harmonique.

3.6.4 Unicité de la solution du problème de Dirichlet

Le problème de Dirichlet (pour l'opérateur laplacien Δ) demande la résolution de l'équation

$$\Delta u = f \quad \text{dans } \Omega$$

avec

$$u = \varphi \quad \text{sur } \operatorname{Fr}\Omega$$

où Ω est une partie ouverte de $\mathbb{R}^N$, $f \in C(\Omega)$, $\varphi \in C(\operatorname{Fr}\Omega)$; Ω, f, φ étant donnés, on cherche une fonction $u : \overline{\Omega} \to \mathbb{R}$ qui est continue dans $\overline{\Omega}$, qui est de classe C^2 dans Ω, qui satisfait l'***équation aux dérivées partielles*** $\Delta u = f$ dans Ω et qui remplit la ***condition au bord*** (dite ***de Dirichlet***) $u = \varphi$ sur $\operatorname{Fr}\Omega$.

Si Ω est une partie ouverte, bornée, non vide alors le problème de Dirichlet précédent possède au plus une solution.

En effet, si u_1, u_2 sont deux solutions du problème alors $u = u_1 - u_2$ est telle que $\Delta u = 0$ dans Ω avec $u = 0$ sur $\operatorname{Fr}\Omega$; du corollaire 3.6.2 on conclut que $u = 0$ dans $\overline{\Omega}$ ce qui donne l'égalité $u_1 = u_2$ dans $\overline{\Omega}$. ♦

Il faut souligner que si l'unicité de la solution du problème de Dirichlet est facile à démontrer, il en va tout autrement pour son existence. On verra (vol. 3) qu'il nous faudra faire quelques hypothèses sur la nature de $\operatorname{Fr}\Omega$ pour pouvoir garantir l'existence d'une solution ; ces hypothèses sont tellement générales qu'elles couvrent tous les cas rencontrés en pratique. Néanmoins, il existe

des domaines bornés Ω pour lesquels le problème de Dirichlet ne possède pas toujours une solution. De plus, si Ω n'est pas borné, l'unicité aussi fait défaut (Ex. 8, sect. 3.11).

3.7 FONCTIONS CONVEXES

Dans le paragraphe 1.9.4 on a vu qu'une fonction convexe définie dans un ensemble convexe ouvert est automatiquement continue. Dans cette section, nous donnons un critère pour vérifier qu'une fonction est convexe.

3.7.1 Proposition

Soit $f : \Omega \to \mathbb{R}$ une fonction de classe C^2 où Ω est un ensemble ouvert et convexe dans $\mathbb{R}^N$. Alors f est convexe si et seulement si sa matrice hessienne est de type positif partout dans Ω.

DÉMONSTRATION

Soit $H(x)$ la matrice hessienne de f au point $x \in \Omega$. Si a, b sont deux points de Ω et $0 < t < 1$, posons

$$c = ta + (1-t)b\,, \quad F = tf(a) + (1-t)f(b) - f(c)\,.$$

Du développement de Taylor d'ordre 2, on a

$$f(a) - f(c) = (a-c)\cdot\nabla f(c) + H(\theta)(a-c)\cdot(a-c)$$
$$f(b) - f(c) = (b-c)\cdot\nabla f(c) + H(\theta')(b-c)\cdot(b-c)$$

où θ, θ' sont deux points appropriés dans les segments $]\,a\,,\,c\,[$, $]\,c\,,\,b\,[$ respectivement. En observant que

$$F = t\{f(a) - f(c)\} + (1-t)\{f(b) - f(c)\}$$
$$t(a-c) + (1-t)(b-c) = 0\,,$$

on obtient

$$F = tH(\theta)(a-c)\cdot(a-c) + (1-t)H(\theta')(b-c)\cdot(b-c)\,. \tag{3.32}$$

De (3.32) on voit que si $H(x)$ est de type positif pour tout x (ce qui veut dire que $H(x)\xi\cdot\xi \geqslant 0$ pour tout ξ dans $\mathbb{R}^N$), alors $F \geqslant 0$; ceci démontre que f est convexe si $H(x)$ est de type positif pour tout $x \in \Omega$.

Supposons maintenant que f est convexe ; il faut voir que $H(x)$ est de type positif pour tout $x \in \Omega$. Sinon, il existe un point $a \in \Omega$ et un $\xi \in \mathbb{R}^N$ tel que

$$H(a)\xi\cdot\xi < 0\,;$$

par la continuité de H on aura que pour $r > 0$, suffisamment petit,

$$H(x)\xi \cdot \xi < 0\,, \quad x \in B(a;r)\,,$$

donc

$$H(x)s\xi \cdot s\xi = s^2 H(x)\xi \cdot \xi < 0 \quad \text{pour } s \in \mathbb{R}^*\,,\ x \in B(a;r)\,.$$

Choisissons un $b \in B(a;r)$ tel que $b-a$ est un multiple non nul de ξ ; si $0 < t < 1$ et $c = ta + (1-t)b$, on aura alors que $(a-c)$ et $(b-c)$ seront aussi des multiples non nuls de ξ d'où

$$H(\theta)(a-c)\cdot(a-c) < 0\,, \quad H(\theta')(b-c)\cdot(b-c) < 0$$

pour θ, θ' quelconques dans $B(a;r)$. De (3.32) on conclut que F est alors < 0 ce qui contredit la convexité de f. ♦

3.7.2 Remarques

De la démonstration on voit que si $H(x)$ est de type strictement positif pour tout $x \in \Omega$, alors f est ***strictement convexe*** dans Ω ce qui veut dire que la quantité F dans la démonstration de la proposition 3.7.1 est > 0 pour tout choix de a, b dans Ω et de t dans $]\,0\,,\,1\,[$. Mais f peut être strictement convexe sans que $H(x)$ soit partout de type strictement positif comme montre l'exemple dans $\mathbb{R}$ de $f(x) = x^4$ dont la hessienne est la dérivée seconde $f''(x) = 12\,x^2$ qui s'annule à $x = 0$; le fait que f est strictement convexe est une conséquence du fait que f' et strictement croissante (Ex. 5, § 3.7.9).

Il est bien clair que de la proposition 3.7.1 découle un critère de concavité pour les fonctions f de classe C^2 définie dans une partie ouverte et convexe de $\mathbb{R}^N$: il est nécessaire et suffisant pour cela que H soit partout de type négatif.

La condition qu'une fonction réelle f de N variables réelles soit convexe est en fait équivalente à la condition de convexité de la restriction de f sur tout segment contenu dans le domaine de définition de f ; ainsi la plupart des propriétés de f se déduisent de celles de la fonctions d'une variable réelle φ définie par

$$\varphi(t) = f\big(a + t(b-a)\big)\,, \quad 0 \leqslant t \leqslant 1\,,$$

a, b étant deux points dans le domaine de définition de f.

Nous laissons comme un exercice la déduction de la proposition 3.7.1 à partir de la validité de celle-ci pour les fonctions d'une variable réelle comme φ. Nous illustrons cette méthode dans la démonstration de la proposition suivante.

3.7.3 Proposition

Soit $f : \Omega \to \mathbb{R}$ une fonction dérivable où Ω est un ensemble ouvert et convexe dans $\mathbb{R}^N$. Alors f est convexe si et seulement si, pour tout $c \in \Omega$,

$$f(x) \geqslant f(c) + (x-c)\cdot \nabla f(c)\,, \quad x \in \Omega\,. \tag{3.33}$$

DÉMONSTRATION

Etablissons d'abord l'analogue de (3.33) pour une fonction réelle convexe, dérivable φ définie dans un intervalle J de $\mathbb{R}$. La première remarque à faire est que la pente, pour un c fixe dans J,

$$m(c,x) = \frac{\varphi(x) - \varphi(c)}{x - c} = m(x,c)$$

est une fonction croissante de $x \in J$. En effet, soit $a < c < b$, a, b dans J ; puisque

$$c = \frac{b-c}{b-a}\, a + \frac{c-a}{b-a}\, b\,,$$

la convexité de φ entraîne

$$\varphi(c) \leqslant \frac{b-c}{b-a}\, \varphi(a) + \frac{c-a}{b-a}\, \varphi(b)$$

d'où un petit calcul montre que

$$m(a,c) \leqslant m(a,b) \leqslant m(b,c)\,,$$

ce qui montre que $x \mapsto m(x,c)$ est croissante. Donc, si $x > c$, on a pour $\varepsilon > 0$ assez petit,

$$m(x,c) \geqslant m(c+\varepsilon, c)\,,$$

d'où, en laissant $\varepsilon \to 0$, on aura

$$m(x,c) \geqslant \varphi'(c)\,,$$

ce qui équivaut à

$$\varphi(x) \geqslant \varphi(c) + (x-c)\varphi'(c)\,.$$

On obtient la même inégalité pour $x < c$ en raisonnant sur

$$m(x,c) \leqslant m(c-\varepsilon, c)\,.$$

Maintenant passons au cas de f, fonction de N variables réelles. Si $c \in \Omega$, $x \in \Omega$, posons

$$\varphi(t) = f(c + t\xi)\,, \quad \xi = (x - c)\,;$$

il est clair que

$$\varphi(0) = f(c)\,, \quad \varphi(1) = f(x)\,,$$

et que φ est convexe et dérivable dans un intervalle ouvert contenant $[0\,,\,1]$; aussi

$$\varphi'(t) = \xi \cdot \nabla f(c + t\xi)\,, \quad \varphi'(0) = \xi \cdot \nabla f(c)\,.$$

De l'inégalité établie pour la fonction convexe d'une variable φ, on aura

$$\varphi(t) \geqslant \varphi(0) + \varphi'(0)t\,,$$

ce qui donne pour $t = 1$

$$f(x) \geqslant f(c) + (x - c) \cdot \nabla f(c) .$$

Ainsi (3.33) est établie pour f convexe et dérivable.

Inversement, si f satisfait (3.33) démontrons que f est convexe. Prenons a, b dans Ω et définissons c, F comme dans la démonstration de la proposition 3.7.1 ; de (3.33) appliquée à $x = a$ et $x = b$ on aura

$$F \geqslant t(a - c) \cdot \nabla f(c) + (1 - t)(b - c) \cdot \nabla f(c) = 0$$

car $t(a - c) + (1 - t)(b - c) = 0$; ceci établit que f est convexe. ♦

3.7.4 Remarques

L'interprétation géométrique de l'inégalité (3.33) est intéressante à souligner ici. Dans l'espace $\mathbb{R}^{N+1}$, décrivons les points comme (x, y), $x \in \mathbb{R}^N$, $y \in \mathbb{R}$; alors si $\nu \in \mathbb{R}^N$, l'équation

$$y = f(c) + (x - c) \cdot \nu$$

représente un hyperplan contenant le point $\big(c, f(c)\big)$; pour $N = 1$, c'est l'équation d'une droite de pente ν, passant par le point $\big(c, f(c)\big)$. L'inégalité (3.33) dit que l'hypersurface de l'équation

$$y = f(x) , \quad x \in \Omega ,$$

est *au-dessus* de l'hyperplan tangent au point $\big(c, f(c)\big)$ pour lequel, par définition, ν ci-dessus est égal à $\nabla f(c)$. L'interprétation géométrique pour $N = 1$ est indiquée dans la figure 3.5.

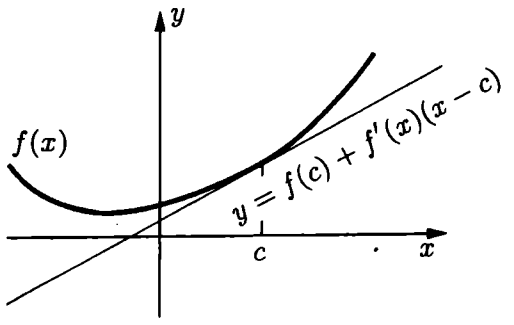


Fig. 3.5

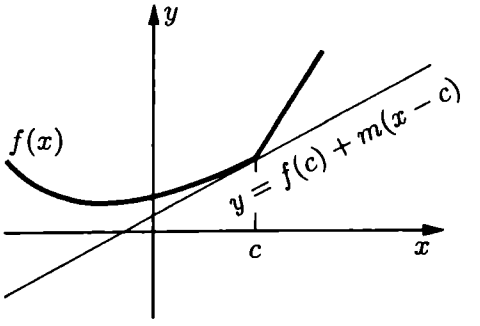


Fig. 3.6

Notons que même si f n'est pas dérivable, une inégalité du type (3.33) caractérise les fonctions convexes. Si f est une fonction convexe d'une variable réelle définie dans un intervalle ouvert J, alors le raisonnement donné dans la démonstration de la proposition 3.7.3 montre que pour tout point $c \in J$, les dérivées à droite et à gauche, $f'_+(c)$, $f'_-(c)$ existent et

$$f'_-(c) \leqslant f'_+(c) ;$$

aussi, $c \mapsto f'_-(c)$, $c \mapsto f'_+(c)$ sont des fonctions croissantes ; de plus, si m est tel que

$$f'_-(c) \leqslant m \leqslant f'_+(c) ,$$

alors la droite d'équation

$$y = f(c) + m(x - c)$$

est au-dessous du graphe $y = f(x)$, c'est-à-dire

$$f(x) \geqslant f(c) + m(x - c) , \quad x \in J , \tag{3.34}$$

comme cela est illustré par la figure 3.6. La démonstration de ces affirmations résulte de la monotonicité de la pente $m(x, c) = m(c, x)$ comme fonction de x, avec

$$m(x, c) = \frac{f(x) - f(c)}{x - c} = m(c, x) ;$$

la preuve de cette monotonicité apparaît dans la démonstration de la proposition 3.7.3.

Pour le cas des fonctions f non nécessairement dérivables, définies dans une partie ouverte, convexe Ω de $\mathbb{R}^N$, il subsiste une inégalité du type (3.33) où l'on remplace $\nabla f(c)$ par un vecteur approprié $\widetilde{\nu} \in \mathbb{R}^N$. Ceci est une conséquence d'un résultat général et important concernant les ensembles convexes généraux que nous décrivons maintenant.

Soit A une partie convexe non vide dans un espace $\mathbb{R}^M$: *si* $a \in \operatorname{Fr} A$ *alors il existe un vecteur non nul* $v \in \mathbb{R}^M$ *et un nombre* $\alpha \in \mathbb{R}$ *tel que*

$$a \cdot v = \alpha ; \quad z \cdot v \leqslant \alpha \quad \text{pour tout } z \in A .$$

On décrit cet énoncé en utilisant une terminologie géométrique qui suit : l'ensemble des points $z \in \mathbb{R}^M$ tels que $z \cdot v = \alpha$ décrit un hyperplan dans $\mathbb{R}^M$; les points z tels que $z \cdot v \leqslant \alpha$ détermine un des deux demi-espaces de $\mathbb{R}^M$ définis par cet hyperplan, l'autre étant $z \cdot v > \alpha$. On peut formuler ainsi l'énoncé précédent comme suit : si $a \in \operatorname{Fr} A$, alors il existe un hyperplan π passant par a tel que l'ensemble A est contenu dans un des demi-espaces définis par cet hyperplan. Un tel hyperplan π passant par a s'appelle un ***hyperplan d'appui*** de A au point a. Avec cette terminologie, l'énoncé dit simplement que *si* A *est convexe et* $a \in \operatorname{Fr} A$, *alors il existe un hyperplan d'appui de* A *au point* a.

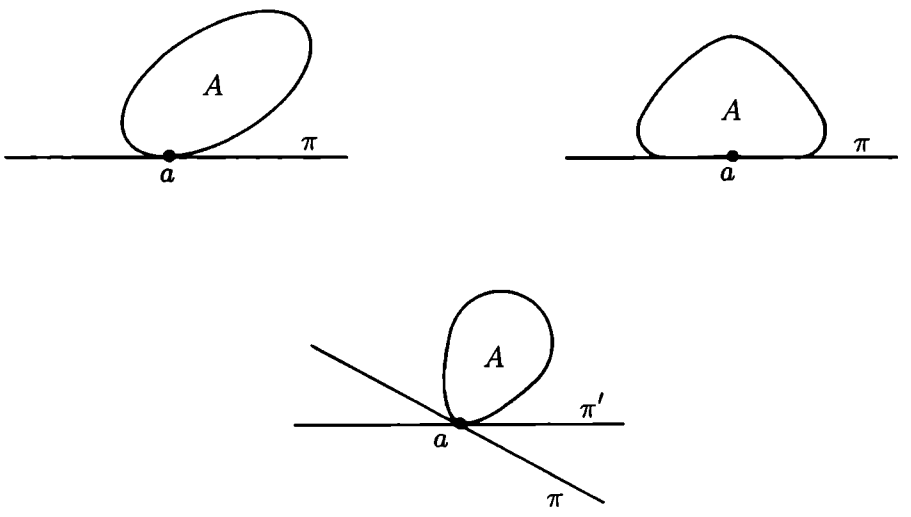


Fig. 3.7

La figure 3.7 montre les différentes possibilités géométriques : l'hyperplan d'appui π peut être uniquement déterminé par a ou non ; même s'il est uniquement déterminé, il peut contenir des points de A autres que a.

Démontrons maintenant une inégalité du type (3.33) à partir de l'énoncé précédent.

Si $f : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^N$, est donnée, on forme l'ensemble appelé l'***épigraphe*** de f :

$$A = A(f) = \{(x, y) : x \in \Omega,\ y \geqslant f(x)\} \subset \mathbb{R}^{N+1}\,.$$

On voit facilement que, si Ω est convexe, la fonction f est convexe si et seulement si l'ensemble A est convexe (Ex. 17, sect. 1.10). Supposons que Ω est ouvert et convexe, et soit $c \in \Omega$; le point $a = (c, f(c))$ appartient à Fr A. L'hyperplan d'appui de A au point a est déterminé par un vecteur $v = (\nu, s) \in \mathbb{R}^{N+1}$, $\nu \in \mathbb{R}^N$, $s \in \mathbb{R}$ et un nombre $\alpha \in \mathbb{R}$ tel que si $z = (x, y) \in \mathbb{R}^{N+1}$, avec $x \in \mathbb{R}^N$, $y \in \mathbb{R}$, on aura

$$a \cdot v = c \cdot \nu + f(c)s = \alpha\,, \quad z \cdot v = x \cdot \nu + ys \leqslant \alpha \text{ si } z \in A\,.$$

On voit aisément que $s < 0$ ($s > 0$ est impossible en laissant $y \to \infty$, $s = 0$ est impossible car $c \in \Omega$ et Ω est ouvert) d'où, de ce qui précède,

$$x \cdot \nu + ys \leqslant c \cdot \nu + f(c)s$$

dès que $y \geqslant f(x)$ et $x \in A$; ainsi, en prenant en compte que $s < 0$, on aura

$$y \geqslant f(c) + (x - c) \cdot \frac{-\nu}{s}\,,$$

c'est-à-dire, en laissant $y \to f(x)$,

$$f(x) \geqslant f(c) + (x - c) \cdot \widetilde{\nu}\,, \quad x \in \Omega\,,$$

où $\widetilde{\nu} = -\nu/s$.

Les fonctions convexes sont à la base de nombreuses inégalités importantes d'analyse. Nous en donnons quelques-unes maintenant.

3.7.5 Inégalité de Jensen

Soit $f : \Omega \to \mathbb{R}$ une fonction convexe définie dans un ensemble convexe $\Omega \subset \mathbb{R}^N$; alors

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leqslant \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n) \tag{3.35}$$

pour tout choix de $x_1, \ldots, x_n$ dans Ω et de $\alpha_1, \ldots, \alpha_n$ dans $[0, 1]$ tels que $\alpha_1 + \cdots + \alpha_n = 1$.

DÉMONSTRATION

Pour $n = 2$, (3.35) est la définition de la convexité de f ; le cas général est établi par récurrence ; pour passer de n à $n + 1$ on pourra prendre

$$a = x_1\,, \quad b = \frac{\alpha_2 x_2 + \cdots + \alpha_{n+1} x_{n+1}}{1 - \alpha_1}$$

en supposant, sans perte de généralité, que $0 < \alpha_1 < 1$; alors, b est une combinaison convexe des n vecteurs $x_2, \ldots, x_{n+1}$ et

$$\alpha_1 a + (1 - \alpha_1) b = \alpha_1 x_1 + \cdots + \alpha_{n+1} x_{n+1}$$

d'où, selon l'hypothèse de récurrence et la convexité de f, on a

$$f(b) \leqslant \frac{\alpha_2}{1 - \alpha_1} f(x_2) + \cdots + \frac{\alpha_{n+1}}{1 - \alpha_1} f(x_{n+1})$$

et

$$f\big(\alpha_1 a + (1 - \alpha_1) b\big) \leqslant \alpha_1 f(a) + (1 - \alpha_1) f(b)\,,$$

ce qui donne (3.35) avec $n + 1$ en place de n. ♦

3.7.6 Inégalité de Minkowski

Si $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$, posons

$$\|x\|_p = \left(\sum_{i=1}^{N} |x_i|^p \right)^{1/p}\,, \quad 0 < p < \infty\,.$$

On a, pour x, y dans $\mathbb{R}^N$ et $p \geqslant 1$

$$\|x + y\|_p \leqslant \|x\|_p + \|y\|_p\,. \tag{3.36}$$

DÉMONSTRATION

Si $x = (x_1, \ldots, x_N)$, $y = (y_1, \ldots, y_N)$, il suffit de démontrer (3.36) pour $x_i \geqslant 0$, $y_i \geqslant 0$, $x_i + y_i > 0$, $i = 1, \ldots, N$. Posons

$$f(t) = \|tx + (1-t)y\|_p, \quad 0 \leqslant t \leqslant 1,$$

et vérifions que f est une fonction convexe si $p \geqslant 1$. Un calcul direct montre que, si $0 < t < 1$, on a

$$f''(t) = (p-1)\left(z_1^p + \cdots + z_N^p\right)^{-2+1/p} \cdot E$$

où $z_j = z_j(t) = tx_j + (1-t)y_j$, $j = 1, \ldots, N$ et

$$E = \left(\sum_{j=1}^{N} z_j^p\right)\left(\sum_{j=1}^{N}(x_j - y_j)^2 z_j^{p-2}\right) - \left\{\sum_{j-1}^{N}(x_j - y_j)z_j^{p-1}\right\}^2.$$

En écrivant

$$(x_j - y_j)z_j^{p-1} = (x_j - y_j)z_j^{p/2-1} z_j^{p/2},$$

on obtient, de l'inégalité de Cauchy-Schwarz,

$$E \geqslant 0$$

d'où $f''(t) \geqslant 0$, $0 < t < 1$, si $p \geqslant 1$. Ceci établit la convexité de f dans $[0, 1]$ si $p \geqslant 1$ (car f est continue dans cet intervalle fermé ; Ex. 1, § 3.7.9). Ainsi

$$f\left(\frac{1}{2}\right) = \frac{1}{2}\{f(0) + f(1)\},$$

ce qui donne (3.36) compte tenu du fait que

$$f(1) = \|x\|_p, \quad f(0) = \|y\|_p, \quad f\left(\frac{1}{2}\right) = \frac{1}{2}\|x + y\|_p. \qquad \blacklozenge$$

3.7.7 Remarques

L'inégalité du paragraphe 3.7.6 montre que

$$x \longmapsto \|x\|_p, \quad x \in \mathbb{R}^N,$$

est une norme dans $\mathbb{R}^N$ si $p \geqslant 1$; le cas $p = 2$ est la norme euclidienne déjà connue ; le cas $p = 1$ de (3.36) est immédiat ; on verra que, si $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$,

$$\lim_{p \to \infty} \|x\|_p = \max_i |x_i| = |x|$$

ce qui justifie l'écriture souvent utilisée (Ex. 14, § 3.7.9) :

$$|x| = \|x\|_\infty.$$

La démonstration du paragraphe 3.7.6 montre que si $x_i \geqslant 0$, $y_i \geqslant 0$, $i = 1, \ldots, N$, alors f est concave si $0 < p < 1$ car alors $f''(t) \leqslant 0$, $0 < t < 1$; on en déduit que si $x = (x_1, \ldots, x_N)$, $y = (y_1, \ldots, y_N)$, $x_i \geqslant 0$, $y_i \geqslant 0$, $i = 1, \ldots, N$ et $0 < p \leqslant 1$, on aura

$$\|x + y\|_p \geqslant \|x\|_p + \|y\|_p . \tag{3.37}$$

3.7.8 Inégalité de Hölder

Si $x = (x_1, \ldots, x_N)$, $y = (y_1, \ldots, y_N)$ *sont dans* $\mathbb{R}^N$ *et* $p > 1$, $q > 1$ *sont tels que* $p^{-1} + q^{-1} = 1$ *alors*

$$\left| \sum_{i=1}^{N} x_i y_i \right| \leqslant \|x\|_p \cdot \|y\|_q . \tag{3.38}$$

DÉMONSTRATION

Il suffit d'établir (3.38) pour $x_i > 0$, $y_i > 0$, $i = 1, \ldots, N$; de plus, on peut supposer que $\|x\|_p = 1$ car, autrement, on passera de x_i à $x_i / \|x\|_p$, $i = 1, \ldots, N$. L'inégalité (3.38) est maintenant une conséquence de la convexité de la fonction f donnée par

$$f(t) = t^q , \quad t > 0 \quad (q > 1) .$$

En effet, compte tenu de $\sum_i x_i^p = 1$ et de la convexité de f, on a

$$\left(\sum_{i=1}^{N} x_i y_i \right)^q = \left(\sum_{i=1}^{N} x_i^p \cdot y_i x_i^{-p+1} \right)^q$$

$$\leqslant \sum_{i=1}^{N} x_i^p \cdot y_i^q x_i^{-pq+q} = \sum_{i=1}^{N} y_i^q$$

car $p + q - pq = 0$. Ceci établit (3.38). ♦

Nombreuses autres inégalités importantes d'analyse mathématique sont des conséquences des inégalités précédentes ou de la convexité d'une fonction appropriée. On en trouvera plusieurs parmi les exercices qui suivent.

3.7.9 Exercices

[1] Soit Ω une partie convexe de $\mathbb{R}^N$; démontrer que $\overline{\Omega}$ est convexe. Si $f : \overline{\Omega} \to \mathbb{R}$ est continue et $f|_{\Omega}$ est convexe, alors f est convexe.

[2] Soit $f : \Omega \to \mathbb{R}$ où Ω est une partie convexe de $\mathbb{R}^N$; vérifier que f est convexe si et seulement si la restriction de f à chaque segment de Ω est convexe.

3 Soit $f : \Omega \to \mathbb{R}$ une fonction convexe définie dans une partie convexe Ω de $\mathbb{R}^N$; si $g : J \to \mathbb{R}$ est une fonction croissante et convexe définie dans un intervalle $J \supset f(\Omega)$, alors la fonction $g \circ f : \Omega \to \mathbb{R}$ est convexe.

4 Soit $f : \Omega \to \mathbb{R}$ une fonction convexe dérivable définie dans une partie ouverte convexe Ω de $\mathbb{R}^N$. Si $a \in \Omega$ est un point critique de f, démontrer que c'est un point de minimum absolu pour f.

5 Soit f une fonction réelle dérivable définie dans un intervalle ouvert de $\mathbb{R}$. Démontrer que f est convexe si f' est croissante et que f est strictement convexe si f' est strictement croissante.

6 Vérifier que les fonctions réelles suivantes sont convexes :

(i) $f(x) = x^p$, $x > 0$ ($p \geqslant 1$ ou $p < 0$) ;
(ii) $f(x) = x \ln x$, $x > 0$;
(iii) $f(x) = \mathrm{e}^x$, $x \in \mathbb{R}$;
(iv) $f(x) = x/(1-x)$, $0 < x < 1$.

7 En utilisant la convexité de $f(x) = \mathrm{e}^x$, $x \in \mathbb{R}$, démontrer que

$$\exp(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leqslant \alpha_1 \mathrm{e}^{x_1} + \cdots + \alpha_n \mathrm{e}^{x_n}$$

si $x_i \in \mathbb{R}$, $\alpha_i \geqslant 0$, $i = 1, \ldots, n$, $\sum_{i=1}^n \alpha_i = 1$. En posant $a_i = \mathrm{e}^{x_i}$ on obtient l'inégalité entre les moyennes arithmétiques et géométriques pondérées :

$$a_1^{\alpha_1} a_2^{\alpha_2} \cdots a_n^{\alpha_n} \leqslant (\alpha_1 a_1 + \cdots + \alpha_n a_n)$$

si les $a_i > 0$ et les α_i comme ci-dessus ; si les $\alpha_i > 0$ alors il y a inégalité stricte ci-dessus sauf si $a_1 = a_2 = \ldots = a_n$.

8 (Inégalité de la théorie d'information) Si $p_i \geqslant 0$, $i = 1, \ldots, n$, $\sum_{i=1}^n p_i = 1$, posons

$$I(p_1, \ldots, p_n) = -\sum_{i=1}^n p_i \ln p_i \quad (\text{avec } 0 \ln 0 = 0).$$

Démontrer que

$$I(p_1, \ldots, p_n) \leqslant \ln n = I\left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$$

avec égalité seulement si $p_i = \ldots = p_n = 1/n$. (Utiliser la convexité de $x \mapsto x \ln x$, $x > 0$; $I(p_1, \ldots, p_n)$ s'appelle l'information de Shannon de la répartition de probabilité $\{p_i\}$; ainsi, l'inégalité dit que I est maximal pour la répartition uniforme $\{1/n, \ldots, 1/n\}$.)

9 Si $\alpha_i \geqslant 0$, $x_i \geqslant 0$, $i = 1, \ldots, n$, $\sum_{i=1}^n \alpha_i = 1$, démontrer que

$$\left(\sum_{i=1}^n \alpha_i x_i^r\right)^{1/r} \leqslant \left(\sum_{i=1}^n \alpha_i x_i^s\right)^{1/s} \quad \text{si } 0 < r < s\,.$$

(Utiliser la convexité de $x \to x^p$, $p = s/r$, $x > 0$.)

10 Si $0 \leqslant x_i < 1$, $i = 1, \ldots, n$ et $S = \sum_{i=1}^n x_i$ démontrer que

$$\sum_{i=1}^n \frac{x_i}{1 - x_i} \geqslant \frac{nS}{n - S}$$

avec égalité seulement si les x_i sont égaux.

11 Vérifier que $f(x) = \left(1 + x^r\right)^{1/r}$, $x > 0$, définit une fonction convexe f si $r > 1$. Conclure que

$$f\left(\frac{\sum_{i=1}^n \alpha_i x_i}{\sum_{i=1}^n \alpha_i}\right) \leqslant \frac{\sum_{i=1}^n \alpha_i f(x_i)}{\sum_{i=1}^n \alpha_i}\,,$$

si $x_i \geqslant 0$, $\alpha_i \geqslant 0$, $i = 1, \ldots, n$, $\sum_{i=1}^n \alpha_i > 0$. En prenant $x_i = b_i/a_i$, $\alpha_i = a_i > 0$, $b_i \geqslant 0$ démontrer que

$$\left\{\left(\sum_{i=1}^n a_i\right)^r + \left(\sum_{i=1}^n b_i\right)^r\right\}^{1/r} \leqslant \sum_{i=1}^n \left(a_i^r + b_i^r\right)^{1/r}$$

ce qui est l'inégalité de Minkowski d'indice $r > 1$ dans $\mathbb{R}^2$ appliquée aux vecteurs $(a_1, b_1), \ldots, (a_n, b_n)$.

12 Soit Ω une partie convexe de $\mathbb{R}^N$. Si f_i, $1 \leqslant i \leqslant n$, sont des fonctions convexes dans Ω et $c_i \geqslant 0$, $1 \leqslant i \leqslant n$, alors $f = \sum_{i=1}^n c_i f_i$ est une fonction convexe dans Ω. Si $\{f_\alpha\}$ est une famille quelconque de fonctions convexes dans Ω alors $f(x) = \sup_\alpha f_\alpha(x)$ définit une fonction convexe dans Ω, pourvu que $f(x) < \infty$ pour tout $x \in \Omega$. En particulier, si les f_α sont des fonctions affines alors $f = sup_\alpha f_\alpha$ définit une fonction convexe dans Ω, pourvu qu'elle soit finie partout. Inversement, toute fonction convexe f dans Ω est de la forme $\sup_\alpha f_\alpha$ où les f_α sont les fonctions affines, c'est-à-dire de la forme $f_\alpha(x) = a_\alpha \cdot x + b_\alpha$, a_α, b_α étant dans $\mathbb{R}^N$. Pour f dérivable ceci est une conséquence de la proposition 3.7.3 ; voir le paragraphe 3.7.4 pour le cas général.

13 La limite (finie) d'une suite de fonctions convexes dans une partie convexe de $\mathbb{R}^N$ définit une fonction convexe.

14 Si $x \in \mathbb{R}^N$ démontrer que

$$\lim_{p\to\infty} \|x\|_p = \max_i |x_i|$$

où $x = (x_1, \ldots, x_N)$.

3.8 LE LAPLACIEN EN DIFFÉRENTES COORDONNÉES

3.8.1 Introduction

Nous expliquons d'abord la position du problème général. Soit Ω une partie ouverte de $\mathbb{R}^N$; si $\widetilde{\Omega}$ est une partie ouverte de $\mathbb{R}^N$ (imaginée dans un «autre» espace $\mathbb{R}^N$ dont les points seront notés par u) et $T : \widetilde{\Omega} \to \Omega$ est une bijection, alors à chaque point x de Ω est associé un unique point u de $\widetilde{\Omega}$ tel que $T(u) = x$; on dit souvent que $u = (u_1, \ldots, u_N)$ sont les ***coordonnées curvilignes*** du point x définies par T. En général cette terminologie est utilisée seulement si T et T^{-1} sont au moins de la classe C^1 ; on dit alors que T est un C^1-***difféomorphisme*** entre Ω et $\widetilde{\Omega}$. Si $f : \Omega \to \mathbb{R}$ alors $g = f \circ T$ est une fonction réelle dans $\widetilde{\Omega}$; si f est de classe C^2 et T est de classe C^2, alors g aussi est de classe C^2 et la question qui se pose est de calculer $\Delta f(x)$ en termes des dérivées partielles de g calculées au point u tel que $Tu = x$. Dans cette section nous donnerons ces calculs pour certaines transformations T importantes associées aux coordonnées curvilignes souvent utilisées.

Dans la suite nous prenons un point de vue moins strict qui est également habituel. Nous considérons deux fonctions f et g reliées entre elles par la formule

$$g(u) = f(Tu)\,,$$

où T est de classe C^2 dans un ensemble ouvert A et f est de classe C^2 dans un ensemble ouvert B contenant $T(A)$. Nous ne spécifierons pas les ensembles A et B exactement, notre objectif étant seulement d'obtenir une formule pour $\Delta f(x)$ en termes des dérivées partielles de g calculées à un point u tel que $Tu = x$. Notons que si f, T sont comme décrits ci-dessus, alors g sera aussi de classe C^2 dans l'ensemble ouvert A. Dans beaucoup de discussions dans les sciences appliquées on va même jusqu'à écrire $f(u)$ au lieu de $g(u)$, la transformation Tu étant admise tacitement dans les calculs.

3.8.2 Le laplacien dans $\mathbb{R}^2$ en coordonnées polaires

Ici

$$T(r,\theta) = (r\cos\theta\,,\, r\sin\theta)\,,$$

ce que l'on écrit habituellement (sect. 3.5) comme

$$x = r\cos\theta\,, \quad y = r\sin\theta\,.$$

Soit f une fonction réelle de classe C^2 ; posons

$$g(r,\theta) = f(r\cos\theta\,,\, r\sin\theta) = f(x,y)\,.$$

Alors

$$\triangle f = \frac{\partial^2 g}{\partial r^2} + \frac{1}{r}\frac{\partial g}{\partial r} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} \tag{3.39}$$

où le côté gauche est calculé en un point (x,y) et le côté droit au point correspondant (r,θ). Cette dernière précision sera souvent omise dans la suite.

La formule (3.39) s'obtient par les calculs suivants basés sur la dérivation des fonctions composées. On a

$$g'_r = f'_x\cos\theta + f'_y\sin\theta\,, \quad g'_\theta = r(-f'_x\sin\theta + f'_y\cos\theta)$$

d'où

$$g''_{rr} = f''_{xx}\cos^2\theta + 2f''_{xy}\sin\theta\cos\theta + f''_{yy}\sin^2\theta$$
$$g''_{\theta\theta} = r^2\left(f''_{xx}\sin^2\theta - 2f''_{xy}\sin\theta\cos\theta + f''_{yy}\sin^2\theta\right) - r(f'_x\cos\theta + f'_y\sin\theta)$$

ce qui donne (3.39).

3.8.3 Le laplacien dans $\mathbb{R}^3$ en coordonnées sphériques

Les coordonnées sphériques dans $\mathbb{R}^3$ sont définies par les formules suivantes : $T(r,\theta,\varphi) = (x,y,z)$ où

$$x = r\sin\theta\cos\varphi\,, \quad y = r\sin\theta\sin\varphi\,, \quad z = r\cos\theta \tag{3.40}$$

avec $r > 0$, $0 < \theta < \pi$, $-\pi < \varphi < \pi$. On vérifie que T donne une bijection entre ce dernier ensemble des (r,θ,φ) et l'ensemble de (x,y,z) dans le complémentaire du demi-plan $\{(x,0,z) : x \leqslant 0\,,\ z \in \mathbb{R}\}$.

Notons que $r = r_0 > 0$ définit la sphère de rayon r_0 centré à $(0,0,0)$; $r = r_0$ et $\theta = \theta_0$ définissent une ligne de latitude ou un petit cercle sur cette sphère ; $r = r_0$ et $\varphi = \varphi_0$ définissent une longitude ou un arc de méridien ; l'équation $\theta = \theta_0$ définit un cône simple du sommet $(0,0,0)$ d'équation cartésienne $x^2 + y^2 = (\mathrm{tg}^2\,\theta_0)z^2$ et l'équation $\varphi = \varphi_0$ définit un plan passant par $(0,0,0)$ d'équation cartésienne $(-\sin\varphi_0)x + (\cos\varphi_0)y = 0$.

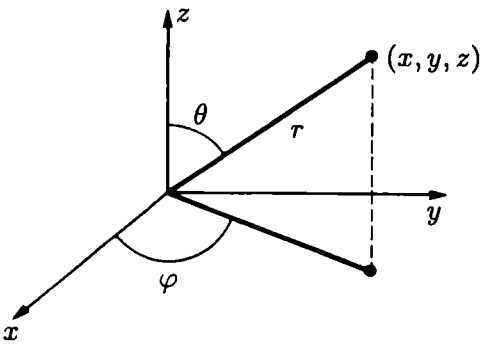


Fig. 3.8

Les nombres r, θ, φ de (3.40) s'appellent les ***coordonnées sphériques*** du point $(x, y, z) \neq (0, 0, 0)$; φ n'est pas défini pour les points $(0, 0, z)$, $z \in \mathbb{R}$; θ s'appelle la ***colatitude*** (la ***latitude*** étant $\pi/2 - \theta$) et φ s'appelle la ***longitude*** du point (x, y, z). Notons les formules analytiques suivantes pour r, θ, φ comme fonctions de (x, y, z) :

$$\left.\begin{aligned} r(x,y,z) &= \left(x^2 + y^2 + z^2\right)^{1/2} \\ \theta(x,y,z) &= \arccos \frac{z}{r} \\ \varphi(x,y,z) &= \text{argument principal de } (x,y) \in \mathbb{R}^2. \end{aligned}\right\} \tag{3.41}$$

On voit que r, θ, φ considérées comme fonctions de (x, y, z) dans le complémentaire de $\{(x, 0, z) : x \leqslant 0\,,\ z \in \mathbb{R}\}$ sont de classe C^∞.

Démontrons que si

$$g(r, \theta, \varphi) = f(x, y, z)$$

où (x, y, z), (r, θ, φ) sont reliés par (3.40) et f est de classe C^2, alors

$$\Delta f = \frac{\partial^2 g}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 g}{\partial \varphi^2} + \frac{2}{r}\frac{\partial g}{\partial r} + \frac{\operatorname{ctg}\theta}{r^2}\frac{\partial g}{\partial \theta} + \cdots \tag{3.42}$$

On va déduire (3.42) par l'application de la formule (3.39) deux fois successivement, en observant que

$$\begin{aligned} f(x,y,z) &= f(\rho\cos\varphi,\ \rho\sin\varphi,\ z) = h(\rho,\varphi,z) \\ &= f(r\sin\theta\cos\varphi,\ r\sin\theta\sin\varphi,\ r\cos\theta) = g(r,\theta,\varphi) \end{aligned}$$

où le passage de f à h se fait par la transformation aux ***coordonnées cylindriques*** ρ, φ, z définies par

$$x = \rho\cos\varphi\,, \quad y = \rho\sin\varphi\,, \quad z = z \tag{3.43}$$

et le passage de h à g se fait par

$$z = r\cos\theta\,, \quad \rho = r\sin\theta\,, \quad \varphi = \varphi\,. \tag{3.44}$$

De (3.39), on obtient

$$\triangle f = h''_{\rho\rho} + \rho^{-2} h''_{\varphi\varphi} + \rho^{-1} h'_\rho + h''_{zz} \tag{3.45}$$

et

$$h''_{\rho\rho} + h''_{zz} = g''_{rr} + r^{-2} g''_{\theta\theta} + r^{-1} g'_r \,. \tag{3.46}$$

Notons que $h''_{\varphi\varphi} = g''_{\varphi\varphi}$; de l'égalité

$$g(r, \theta, \varphi) = h(r \sin\theta,\, \varphi,\, r\cos\theta)\,,$$

on a

$$g'_r = h'_\rho \sin\theta + h'_z \cos\theta\,, \quad g'_\theta = h'_\rho (r\cos\theta) + h'_z(-r\sin\theta)$$

d'où

$$h'_\rho = g'_r \sin\theta + \frac{\cos\theta}{r}\, g'_\theta$$

et

$$\rho^{-1} h'_\rho = r^{-1} g'_r + r^{-2} \operatorname{ctg}\theta\, g'_\theta \tag{3.47}$$

en utilisant $\rho = r\sin\theta$. Mettant ensemble (3.45), (3.46) et (3.47), et compte tenu de (3.44) et $h''_{\varphi\varphi} = g''_{\varphi\varphi}$, on obtient la formule (3.42) voulue.

3.8.4 Inversion, transformée de Kelvin

La transformation T dans $\mathbb{R}^N$ définie par

$$T(x) = \frac{x}{\|x\|^2}\,, \quad x \neq 0$$

s'appelle l'***inversion*** par rapport à la sphère unité ; pour tout $x \neq 0$, les points x, $T(x)$ et 0 sont colinéaires et

$$\big\|T(x)\big\| \cdot \|x\| = 1\,.$$

Notons que T est une bijection de $\mathbb{R}^N \setminus \{0\}$ dans lui-même telle que $T\big(T(x)\big) = x$, d'où

$$T(A) = T^{-1}(A)$$

si A et une partie quelconque de $\mathbb{R}^N \setminus \{0\}$; aussi T est évidemment de classe C^∞ dans son domaine de définition $\mathbb{R}^N \setminus \{0\}$.

Au lieu de considérer $f\big(T(x)\big)$ pour une fonction f donnée, nous allons considérer la fonction g définie par

$$g(x) = \|x\|^{2-N} f\big(T(x)\big) = \|x\|^{2-N} f\left(\frac{x}{\|x\|^2}\right). \tag{3.48}$$

On dit que g est la ***transformée de Kelvin*** de f. On a alors la formule suivante :

$$\triangle g(x) = \|x\|^{-2-N} \triangle f\left(\frac{x}{\|x\|^2}\right), \tag{3.49}$$

pourvu que $f \in C^2(\Omega)$, Ω étant une partie ouverte de $\mathbb{R}^N$ ne contenant pas 0 et $x \in T(\Omega)$.

L'importance de la formule (3.49) est qu'elle entraîne immédiatement ce qui suit : *si f est une fonction harmonique dans Ω, alors sa transformée de Kelvin g est une fonction harmonique dans $T(\Omega)$*. Noter aussi que si $N = 2$ alors $g(x) = f\big(x/\|x\|^2\big)$.

Pour établir (3.49) on fait les calculs suivants esquissés. Pour simplifier l'écriture, nous écrivons f_i, f_{ij}, ... pour $\mathrm{D}_i f$, $\mathrm{D}_{ij} f$, ... et $x = (x_1, \ldots, x_N)$ dont l'écriture sera omise si cela ne cause pas de confusion. On a

$$g = r^{2-N} h\,, \quad h = f \circ T\,, \quad r = \|x\|$$

d'où

$$\begin{aligned} \triangle g &= r^{2-N} \triangle h + 2\nabla\big(r^{2-N}\big) \cdot \nabla h + \triangle r^{2-N} h \\ &= r^{2-N} \triangle h + 2\nabla\big(r^{2-N}\big) \cdot \nabla h \end{aligned} \tag{3.50}$$

(Ex. 8, Ex. 5, §3.2.6). On vérifie que

$$\nabla r^{2-N} = (2-N) r^{-N} x$$

$$\mathrm{D}_i (f \circ T) = r^{-2} f_i \circ T - 2r^{-4} \sum_{j=1}^{N} x_i x_j f_j \circ T$$

d'où, en remplaçant f par f_i dans la dernière formule, on a

$$\begin{aligned} \mathrm{D}_{ii}(f \circ T) = {} & -2r^{-4} x_i f_i \circ T + r^{-2} \left\{ r^{-2} f_{ii} \circ T - 2r^{-4} \sum_j x_i x_j f_{ij} \circ T \right\} \\ & - 2\big(r^{-4} - 4r^{-6} x_i^2\big) \sum_j x_j f_j \circ T \\ & - 2r^{-4} x_i \sum_j x_j \left\{ r^{-2} f_{ij} \circ T - 2r^{-4} x_i \sum_k x_k f_{kj} \circ T \right\} \\ & - 2r^{-4} x_i f_i \circ T\,. \end{aligned}$$

Posons $S = \sum_{i,j} x_i x_j f_{ij} \circ T$; on aura

$$\triangle (f \circ T) = r^{-4} (\triangle f) \circ T + 2(2-N) r^{-4} S$$

et

$$\nabla \dot{(r^{2-N})} \cdot \nabla h = -(2-N) r^{-2-N} S$$

d'où, en vue de (3.50), on obtient la formule (3.49) voulue.

3.9 QUELQUES DÉFINITIONS GÉNÉRALES

3.9.1 Définitions de $C^k(A;\mathbb{R}^M)$

La définition usuelle de la dérivée d'une fonction f en un point x fait intervenir l'expression $f(x+h)-f(x)$ pour h de norme suffisamment petite ; pour que cette expression ait un sens pour tout h avec $\|h\|$ suffisamment petite, il faut que x soit un point intérieur au domaine de définition de f. C'est la raison pour laquelle nous nous sommes restreints aux fonctions définies dans des ensembles ouverts pour discuter de leur dérivabilité. Néanmoins, il nous sera utile d'introduire une terminologie qui nous permettra de parler des dérivées des fonctions définies dans un sous-ensemble quelconque de $\mathbb{R}^N$.

Soit A une partie quelconque de $\mathbb{R}^N$; soit $f : A \to \mathbb{R}^M$ et $x \in A$. S'il existe un voisinage ouvert V de x et une fonction $g : V \to \mathbb{R}^M$ telle que g est dérivable en x et $g(y) = f(y)$ pour y dans $V \cap A$, alors on dira que f est ***dérivable*** en x et on posera $f'(x) = g'(x)$. On peut démontrer que si f est dérivable dans ce sens en tout point x de A, alors il existe un ensemble ouvert V contenant A et une fonction dérivable $g : V \to \mathbb{R}^M$ qui prolonge f (c'est-à-dire $g\big|_A = f$ ce qui veut dire que $g(y) = f(y)$, $y \in A$). On utilisera le même genre de définition pour définir une fonction de classe C^k dans A. Plus précisément, si $f : A \to \mathbb{R}^M$ est une fonction telle qu'il existe un ensemble ouvert V contenant A et une fonction $g : V \to \mathbb{R}^M$ de classe C^k ayant la propriété que $g\big|_A = f$, alors nous dirons que f ***est de classe*** C^k ***dans*** A ; aussi, pour $x \in A$, nous définirons $\mathrm{D}^\alpha f(x)$ comme égal à $\mathrm{D}^\alpha g(x)$ (pourvu évidemment que $|\alpha| \leqslant k$). La famille de fonctions de classe C^k dans A à valeurs dans $\mathbb{R}^M$ sera notée $C^k(A;\mathbb{R}^M)$; ici k est un entier > 0 ou ∞. Notons que $C^0(A;\mathbb{R}^M)$ correspond simplement à la classe de fonctions $f : A \to \mathbb{R}^M$ qui sont continues.

La définition proposée ci-dessus n'est pas très utile si A est une partie trop « mince » de $\mathbb{R}^N$ car les valeurs des quantités $\mathrm{D}^\alpha f(x)$ ne sont pas alors uniquement déterminées (Ex. 1, sect. 3.11). Dans la suite, cette définition sera utilisée seulement dans les cas où $A = \overline{\Omega}$, Ω étant une partie ouverte non vide de $\mathbb{R}^N$; alors les dérivées $\mathrm{D}^\alpha f$ sont uniquement déterminées et notre définition permet de parler des dérivées d'une fonction aux points appartenant à la frontière de son domaine de définition.

Plus généralement, soit A une partie de $\mathbb{R}^N$ telle que $\Omega \subset A \subset \overline{\Omega}$; si $f : A \to \mathbb{R}^M$ est de classe C^k dans A, $k \geqslant 1$, selon la définition donnée ci-dessus alors les quantités $\mathrm{D}^\alpha f(x)$ ($x \in A$, $|\alpha| \leqslant k$) ne dépendent pas du prolongement g utilisé dans la définition. En effet, soient g, h deux prolongements de f définis dans les ensembles ouverts V, W respectivement contenant A ; si $x \in A$, on peut trouver une suite de points $\{x_n\}$ dans Ω, $x_n \neq a$ pour tout n, telle que $x_n \to a$ lorsque $n \to \infty$; puisque $V \cap W$ est une partie ouverte contenant x,

les $x_n \in V \cap W$ pour $n \geqslant n_0$ suffisamment grand ; or, dans l'ensemble ouvert $\Omega \cap V \cap W$, g et h sont égales à f d'où

$$\mathrm{D}^\alpha g(x_n) = \mathrm{D}^\alpha h(x_n)\,, \quad n \geqslant n_0$$

et

$$\mathrm{D}^\alpha f(x) = \lim_{n\to\infty} \mathrm{D}^\alpha g(x_n) = \lim_{n\to\infty} \mathrm{D}^\alpha h(x_n)$$

par la continuité de $\mathrm{D}^\alpha g$, $\mathrm{D}^\alpha h$.

Nous convenons d'écrire $C^k(A)$ au lieu de $C^k(A;\mathbb{R})$.

La définition proposée pour $C^k(A;\mathbb{R}^M)$ est, en fait, une simple convention ; les dérivées d'une fonction f de classe C^k dans A se calcule, selon cette définition, à partir d'une autre fonction g, définie dans une partie ouverte, dont f est une restriction. Le problème important est le suivant : étant donnée une fonction f, possède-t-elle un prolongement g de classe C^k ? Nous ne discuterons pas ce problème difficile ici. A titre d'exemple, la fonction réelle donnée par l'expression

$$f(x,y) = \left(1 - x^2 - y^2\right)^{1/2}$$

est de classe C^∞ dans $\overline{\Omega}_\rho$, si Ω_ρ est le disque de centre $(0,0)$ et de rayon $\rho < 1$, mais elle n'est pas de classe C^1 dans $\overline{\Omega}_1$ car les dérivées partielles de f ne sont pas bornées lorsque (x,y) tend vers un point sur le cercle unité.

3.9.2 Difféomorphismes

Soient A, B deux ensembles, $A \subset \mathbb{R}^N$, $B \subset \mathbb{R}^M$; si $f : A \to B$ est une bijection telle que f et f^{-1} sont de classe C^k, $k \geqslant 1$, on dit que f est un ***C^k-difféomorphisme*** entre A et B ; les ensembles A, B sont dits ***difféomorphes*** (plus précisément ***C^k-difféomorphes***) s'il existe un C^k-difféomorphisme ($k \geqslant 1$) $f : A \to B$.

Notons que deux ensembles difféomorphes sont homéomorphes car un difféomorphisme $f : A \to B$ entre les ensembles A et B concernés est évidemment un homéomorphisme (c'est-à-dire que f est bijective et f ainsi que f^{-1} sont continues). Par contre, deux ensembles peuvent être homéomorphes sans être difféomorphes (Ex. 2, sect. 3.11).

Il faut remarquer qu'une application $f : A \to B$ peut bien être un homéomorphisme de classe C^∞ sans que f soit un difféomorphisme car f^{-1} peut ne pas être de classe C^1. Un exemple classique simple de ce phénomène est fourni par $f : \mathbb{R} \to \mathbb{R}$ donnée par $f(x) = x^3$; f est un homéomorphisme de classe C^∞ mais $f^{-1}(x) = x^{1/3}$ ne définit pas une fonction dérivable au point $x = 0$. En revanche, la proposition suivante nous permet de conclure qu'*un homéomorphisme f de classe C^k ($k \geqslant 1$) entre deux ensembles ouverts de $\mathbb{R}^N$ est un C^k-difféomorphisme si et seulement si $f'(x)$ est inversible pour tout x.*

3.9.3 Proposition

Soit $f : A \to B$ un homéomorphisme entre deux ensembles ouverts A, B de $\mathbb{R}^N$; supposons que f est de classe C^k $(k \geqslant 1)$.

(i) Si $f^{-1} : B \to A$ est dérivable alors $f'(x)$ est inversible pour tout $x \in A$ et f^{-1} est de classe C^k.

(ii) Si $f'(x)$ est inversible pour tout $x \in A$ alors $f^{-1} : B \to A$ est de classe C^k.

DÉMONSTRATION

(i) Posons $g = f^{-1}$; du fait que

$$g\big(f(x)\big) = x\,, \quad x \in A$$

et de la dérivabilité de g, on a

$$g'\big(f(x)\big)f'(x) = I\,, \quad x \in A$$

d'où nous déduisons l'inversibilité de $f'(x)$ pour $x \in A$; de plus, en posant $y = f(x)$,

$$g'(y) = \big[f'\big(g(y)\big)\big]^{-1}\,, \quad y \in B\,. \tag{3.51}$$

Les N^2 éléments de la matrice M donnée par $f'\big(g(y)\big)$ définissent des fonctions continues de $y \in B$; la matrice inverse M^{-1} est donc formée aussi des fonctions continues de $y \in B$ car si $M = [a_{ij}]$, alors

$$M^{-1} = [b_{ij}]\,, \quad b_{ij} = (-1)^{i+j}\frac{A_{ji}}{\det M}\,,$$

A_{ji} étant le déteminant de la matrice de type $(N-1\,,\, N-1)$ obtenue en supprimant la ligne numéro j et la colonne numéro i de M ; ceci montre que les b_{ij} sont des fonctions continues de $y \in B$ d'où on conclut que $g = f^{-1}$ est de classe C^1 car si $g = (g_1, \ldots, g_N)$, $\mathrm{D}_j g_i = b_{ij}$. Si f est de classe C^2, le même raisonnement basé sur la formule (3.51) donne que g est de classe C^2 car alors les b_{ij} sont de classe C^1. On conclut, par récurrence, que si f est de classe C^k alors g aussi est de classe C^k.

(ii) Grâce à (i), il suffit d'établir que $g = f^{-1}$ est dérivable si $f'(x)$ est inversible pour tout $x \in A$. Soit $b \in B$; nous démontrons que si $f(a) = b$ et $M = f'(a)$ alors (k est un élément de $\mathbb{R}^N$ pour le reste de la démonstration) :

$$g(b+k) - g(b) - M^{-1}k = |k|S(k)\,, \quad \lim_{|k|\to 0}\big|S(k)\big| = 0\,.$$

En effet, si $g(b+k) = a + h$, on aura

$$f(a+h) = b + k\,, \quad f(a) = b\,, \quad |h| \to 0 \text{ si } |k| \to 0$$

à cause de la continuité de g ; puisque f est dérivable en a, on aura

$$f(a+h) = f(a) + Mh + |h|R(h)\,, \quad \lim_{|h|\to 0} \big|R(h)\big| = 0$$

d'où

$$\begin{aligned} g(b+k) - g(b) - M^{-1}k &= h - M^{-1}\big\{f(a+h) - f(a)\big\} \\ &= -|h|M^{-1}\big(R(h)\big) = |k|S(k) \end{aligned}$$

avec

$$S(k) = -\frac{|h|}{|k|}\, M^{-1}\big(R(h)\big)\,. \tag{3.52}$$

Notons que la relation

$$k = f(a+h) - f(a) = Mh + |h|R(h)$$

donne

$$|k| \geqslant |Mh| - |h| \cdot \big|R(h)\big|\,.$$

Comme M est inversible, il existe un $c > 0$ tel que $|Mh| \geqslant c|h|$ pour tout h ; ainsi

$$|k| \geqslant |h|\,\big(c - \big|R(h)\big|\big)\ .$$

Puisque $R(h) \to 0$ lorsque $|h| \to 0$, on en déduit qu'il existe un $\varepsilon > 0$ tel que si $|h| < \varepsilon$ alors

$$\frac{|h|}{|k|} \leqslant \frac{2}{c}\ ;$$

ainsi $M^{-1}\big(R(h)\big) \to 0$ lorsque $h \to 0$; on en conclut de (3.52) que $S(k) \to 0$ lorsque $k \to 0$. ♦

3.9.4 Exemples de difféomorphismes

Exemple 1

Une application linéaire inversible $T : \mathbb{R}^N \to \mathbb{R}^N$ est un C^∞-difféomorphisme de $\mathbb{R}^N$ dans $\mathbb{R}^N$. En effet, les composantes de $T(x)$ et $T^{-1}(x)$ sont les fonctions linéaires de $x = (x_1, \ldots, x_N)$, donc de classe C^∞. Si A est une partie quelconque de $\mathbb{R}^N$ et $B = T(A)$, on en déduit que A et B sont C^∞-difféomorphes. Il s'ensuit, par exemple, qu'une boule et un ellipsoïde dans $\mathbb{R}^N$ sont C^∞-difféomorphes.

Exemple 2

La relation de C^k-difféomorphisme entre deux ensembles est une relation d'équivalentce ; en notant cette relation par $A \sim B$, ceci signifie que : $A \sim A$, $A \sim B$ entraîne $B \sim A$ et $A \sim B$, $B \sim C$ entraîne $A \sim C$. Par exemple, si $f : A \to B$, $g : B \to C$ sont des C^k-difféomorphismes alors $h = g \circ f$ est un C^k-difféomorphisme entre A et C.

Exemple 3

Si A_i est C^k-difféomorphe à B_i, $i = 1, \ldots, n$, alors $A = A_1 \times \cdots \times A_n$ et $B = B_1 \times \cdots \times B_n$ sont C^k-difféomorphes. De cette observation, on voit que $]0,1[\times]0,1[$ est C^∞-difféomorphe à $\mathbb{R}^2$ ou à $]0,\infty[\times]0,\infty[$.

Exemple 4

La boule $B(0;1) \subset \mathbb{R}^N$ est C^∞-difféomorphe à $\mathbb{R}^N$. En effet, si

$$f(x) = \left(1 - \|x\|^2\right)^{-1/2} x, \quad \|x\| < 1$$

alors $f : B(0;1) \to \mathbb{R}^N$ est de classe C^∞ ; on vérifie aisément que f est injective ; un petit calcul montre que si

$$g(\xi) = \left(1 + \|\xi\|^2\right)^{-1/2} \xi, \quad \xi \in \mathbb{R}^N$$

alors

$$f\big(g(\xi)\big) = \xi, \quad g\big(f(x)\big) = x$$

ce qui montre la surjectivité de f et que $g = f^{-1}$ est de classe C^∞.

Exemple 5

Notre discussion de coordonnées polaires dans $\mathbb{R}^2$ (sect. 3.5) montre que

$$f(r,\theta) = (r\cos\theta\,,\, r\sin\theta)$$

donne un C^∞-difféomorphisme entre $]0,\infty[\times]-\pi,\pi[$ et $\mathbb{R}^2 \setminus \mathbb{R}_-$.

Exemple 6

Les coordonnées sphériques dans $\mathbb{R}^2$ (§ 3.8.3) donnent l'application

$$f(r,\theta,\varphi) = (r\sin\theta\cos\varphi,\ r\sin\theta\sin\varphi,\ r\cos\theta)$$

qui est un C^∞-difféomorphisme entre $]0,\infty[\times]0,\pi[\times]-\pi,\pi[$ et $\mathbb{R}^3 \setminus D$ où $D = \{(x,0,z) : x \leqslant 0,\ z \in \mathbb{R}\}$.

3.10 QUELQUES REMARQUES SUR LES DÉRIVÉES D'ORDRE SUPÉRIEUR

Soit $f : \Omega \to \mathbb{R}^M$, Ω une partie ouverte de $\mathbb{R}^N$. Si f est dérivable dans Ω alors pour chaque $x \in \Omega$ on obtient $f'(x)$, une application linéaire de $\mathbb{R}^N$ dans $\mathbb{R}^M$. En écrivant $\mathcal{L}(E;F)$ pour l'espace vectoriel de toutes les applications linéaires entre les espaces vectoriels E et F, on aura alors une fonction $f' : \Omega \to \mathcal{L}(\mathbb{R}^N;\mathbb{R}^M)$. On peut alors définir la dérivée de cette application dans l'esprit de la définition de la dérivée donnée dans le chapitre précédent. Cela nous permettra de définir $f''(x)$, pour $x \in \Omega$, comme un élément de

$$\mathcal{L}_2(\mathbb{R}^N;\mathbb{R}^M) = \mathcal{L}\big(\mathbb{R}^N;\, \mathcal{L}(\mathbb{R}^N;\mathbb{R}^M)\big)\,.$$

Ce dernier espace peut être identifié comme l'espace des applications bilinéaires de $\mathbb{R}^N \times \mathbb{R}^N$ dans $\mathbb{R}^M$ en posant pour $u \in \mathcal{L}_2(\mathbb{R}^N; \mathbb{R}^M)$

$$(x, y) \longmapsto u(x)(y) = u(x, y) \in \mathbb{R}^M .$$

Ainsi, on aura l'application

$$f'' : \Omega \to \mathcal{L}_2(\mathbb{R}^N; \mathbb{R}^M) .$$

Si celle-ci est dérivable on aura

$$f^{(3)} : \Omega \to \mathcal{L}_3(\mathbb{R}^N; \mathbb{R}^M) ,$$

où $\mathcal{L}_3(\mathbb{R}^N; \mathbb{R}^M)$ est

$$\mathcal{L}\big(\mathbb{R}^N; \mathcal{L}_2(\mathbb{R}^N; \mathbb{R}^M)\big)$$

ce qui s'identifie avec l'espace des applications trilinéaires entre $\mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N$ et $\mathbb{R}^M$. En posant $\mathcal{L}_k(\mathbb{R}^N; \mathbb{R}^M)$ pour l'espace des applications multilinéaires entre $\mathbb{R}^N \times \cdots \times \mathbb{R}^N$ (k fois) et $\mathbb{R}^M$ on aura une application

$$f^{(k)} : \Omega \to \mathcal{L}_k(\mathbb{R}^N; \mathbb{R}^M)$$

si $f^{(k-1)}$ est dérivable.

Le théorème concernant l'égalité des dérivées mixtes devient dans ce contexte la symétrie de l'application multilinéaire $f^{(k)}(x)$ ($x \in \Omega$) dans le sens que, pour $\xi_1, \ldots, \xi_k$ dans $\mathbb{R}^N$,

$$f^{(k)}(x)(\xi_1, \ldots, \xi_k) = f^{(k)}(x)(\xi_{\sigma(1)}, \ldots, \xi_{\sigma(k)})$$

pour toute permutation σ de $\{1, \ldots, k\}$; en particulier,

$$f''(x)(\xi, \eta) = f''(x)(\eta, \xi) , \quad \xi , \eta \text{ dans } \mathbb{R}^N .$$

La formule de Taylor s'écrit alors comme

$$f(x+h) = f(x) + \frac{1}{1!} f'(x) \cdot h + \frac{1}{2!} f''(x) \cdot h^2 + \cdots + \frac{1}{(k-1)!} f^{(k-1)}(x) \cdot h^{k-1} + \frac{1}{k!} f^{(k)}(x+\theta h) \cdot h^k$$

en écrivant

$$f'(x) \cdot h = f'(x)(h) , \quad f''(x) \cdot h^2 = f''(x)(h, h) ,$$

où f est k-fois dérivable, Ω est tel que le segment $[x , x+h]$ est contenu dans Ω et $0 < \theta < 1$. On notera l'avantage de ce formalisme ; la formule de Taylor s'écrit exactement comme dans la théorie élémentaire des fonctions réelles d'une variable réelle.

Le developpement indiqué ci-dessus n'est qu'une esquisse. Pour le faire complètement, nous devons discuter de la dérivabilité d'une fonction

$f : \Omega \to F$ à valeurs dans un espace vectoriel F quelconque de dimension finie. Cette discussion est en fait assez facile, mais nous ne l'avons pas développé parce qu'elle ne nous sera pas utile dans la suite. Du point de vue de la théorie générale, on aurait dû développer la dérivabilité des fonctions $f : \Omega \to F$ lorsque Ω est une partie «ouverte» d'un espace vectoriel E quelconque de dimension finie et F est un autre espace vectoriel de dimension finie. En fait, dans les développements modernes d'analyse, il est essentiel de généraliser les théories encore au cas des fonctions dont le domaine et les valeurs sont contenus dans des espaces de dimension infinie.

Notons que l'espace $\mathcal{L}_k(\mathbb{R}^N;\mathbb{R})$ s'appelle aussi l'espace des ***tenseurs covariants*** d'ordre k ; l'espace $\mathcal{L}_1(\mathbb{R}^N;\mathbb{R})$ est l'espace de formes linéaires sur $\mathbb{R}^N$ (l'espace des tenseurs covariants d'ordre 1 qui se confond avec $\mathbb{R}^N$). Ainsi, $f^{(k)}(x)$, $x \in \Omega$, est un champ de tenseurs covariants symétriques d'ordre k ; nous n'utilisons pas ce point de vue dans la suite.

3.11 EXERCICES

1 Soit $A = \{(x,0) : x \in \mathbb{R}\} \subset \mathbb{R}^2$ et $f : A \to \mathbb{R}$ donnée par $f(x,0) \equiv 0$. En prenant $g(x,y)$ de la forme $y\varphi(x)$, voir que f possède des prolongements donnant des valeurs différentes pour $D_2 f(x,0)$ pour des choix appropriés de φ.

2 Soient $A =]-1\,,\,1[\subset \mathbb{R}$, $B = \{(0,y) : 0 \leqslant y < 1\} \cup \{(x,0) : 0 \leqslant x < 1\}$, $B \subset \mathbb{R}^2$; démontrer qu'il existe un homéomorphisme $f : A \to B$ mais que A, B ne sont pas difféomorphes.
(Voir que si $f : A \to B$ est un homéomorphisme dérivable, alors $f'(c) = (0,0)$ où c est tel que $f(c) = (0,0)$; si $g = f^{-1}$ est aussi dérivable on aurait $g\big(f(t)\big) = t$ qui donne $g'\big(f(c)\big)f'(c) = 1$, une impossibilité.)

3 Soit $\Omega = \{(x,y) : x > 0\,,\ y > 0\}$; posons $r = \sqrt{x^2+y^2}$ et $f(x,y) = \big(h(r) - h(y)\,,\,y\big)$, $(x,y) \in \Omega$, où $h :]0\,,\,\infty[\to]0\,,\,\infty[$ est une fonction continue, strictement croissante avec $h(t) \to 0$ lorsque $t \to 0$, $h(t) \to \infty$ lorsque $t \to \infty$. Vérifier que f est un homéomorphisme de Ω sur Ω qui n'est pas nécessairement un difféomorphisme ; f est un C^1-difféomorphisme si et seulement si h est de classe C^1 et $h' > 0$. Calculer la matrice de $f'(x,y)$ si h est dérivable. En prenant h dérivable, $h' > 0$ mais non de classe C^1, on aura un homéomorphisme dérivable f tel que f^{-1} est aussi dérivable sans que f soit un C^1-difféomorphisme.

4 Soient u, v deux fonctions harmoniques réelles dans une partie ouverte $\Omega \subset \mathbb{R}^N$. Démontrer que uv est harmonique si et seulement si $\nabla u \cdot \nabla v = 0$. En déduire que si u et u^2 sont harmoniques dans Ω, alors u est une fonction constante.

5 Vérifier que les fonctions définies par les expressions suivantes sont harmoniques :

(i) $u(x) = \|x\|^2 - Nx_1^2$, $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$;
(ii) $u(x) = x_1 \cdot \|x\|^{-N}$, $x = (x_1, \ldots, x_N) \in \mathbb{R}^N \setminus \{(0, \ldots, 0)\}$.

6 Soit u une fonction réelle de classe C^2 dans $\mathbb{R}^N$. Vérifier que :

(i) $\Delta u_t(x) = t^2 \, \Delta u(tx)$, $x \in \mathbb{R}^N$, où $u_t(x) = u(tx)$.
(ii) $\Delta(u \circ T) = \Delta u \circ T$ si $T : \mathbb{R}^N \to \mathbb{R}^N$ est une transformation linéaire orthogonale.

7 Soit $u \in C^\infty(\Omega)$, Ω ouvert dans $\mathbb{R}^N$. Vérifier que si u est harmonique, alors $v = \mathcal{D}^\alpha u$ est harmonique dans Ω et $w(x) = x \cdot \nabla u(x)$ définit une fonction harmonique w dans Ω (on peut démontrer que toute fonction harmonique est de classe C^∞).

8 Soit $\Omega = \{(x, y) : y > 0\} \subset \mathbb{R}^2$. Vérifier que $u(x, y) = y(ax + b)$ définit une fonction harmonique dans Ω pour tout choix des constantes réelles a, b et que $u(x, y) = 0$ si $(x, y) \in \operatorname{Fr} \Omega$. Ceci montre que le théorème d'unicité pour le problème de Dirichlet (§ 3.6.4) est en défaut pour les domaines non bornés.

9 Soit $\Omega = \{x \in \mathbb{R}^N : a < \|x\| < b\}$, $0 < a < b < \infty$ $(N \geqslant 2)$; résoudre le problème de Dirichlet $\Delta u = 0$ dans Ω avec $u(x) = \alpha$ si $\|x\| = a$; $u(x) = \beta$ si $\|x\| = b$. (Utiliser les fonctions harmoniques $c \ln \|x\| + d$ si $N = 2$, $c\|x\|^{2-N} + d$ si $N \geqslant 3$.)

10 Soit $u \in C(\overline{\Omega})$, $\Omega = B(0; 1) \subset \mathbb{R}^N$, $N \geqslant 2$; si u est harmonique dans Ω et $u = \varphi$ sur $\operatorname{Fr} \Omega$, résoudre le problème de Dirichlet (dit d'extérieur) dans $(\overline{\Omega})^c$ avec la même condition au bord $v = \varphi$ sur $\operatorname{Fr} \Omega$ et $\lim_{\|x\| \to \infty} v(x) = u(0)$ ou 0 selon que $N = 2$ ou $N \geqslant 3$; démontrer que la solution v est unique. (Utiliser la transformée de Kelvin, § 3.8.4.)

3.12 COMMENTAIRES

Le point de vue moderne est de considérer les dérivées comme des applications linéaires ; pour les dérivées du premier ordre, ceci est présenté dans le chapitre 2 pour les fonctions du type $\mathbb{R}^N \to \mathbb{R}^M$; pour les dérivées d'ordre supérieur, la section 3.10 donne une esquisse. Un développement détaillé pour les fonctions définies dans les espaces de Banach à valeurs dans les espaces de Banach est donné dans les ouvrages de Dieudonné [13, tome 1] et de Scharwtz [38, tome 2] cités dans les références (voir aussi Cartan [8, tome 1] pour une présentation assez accessible).

Les fonctions harmoniques jouent un rôle central dans beaucoup de théories ; du point de vue de la théorie générale, elles sont les solutions d'une ***équation aux dérivées partielles elliptique*** spéciale ; nous aborderons ce sujet plus généralement dans le volume 3. Sur les fonctions harmoniques, on peut consulter Dautray-Lions [12, tome 2] ; l'ensemble de ces ouvrages est une présentation remarquable d'une vaste gamme de problèmes analytiques importants pour les sciences et les techniques.

Le problème principal d'un grand domaine des mathématiques appelé ***topologie différentielle*** est de déterminer si deux ensembles A, B sont difféomorphes ($A \subset \mathbb{R}^N$, $B \subset \mathbb{R}^M$). La plupart des auteurs se bornent à considérer les C^∞-difféomorphismes ; une exception notable est la petite monographie de Munkres [27] où on trouve un peu plus d'informations sur les questions traitées dans la section 3.9. Un excellent petit livre sur la topologie différentielle est celui de Milnor [26] ; en une soixantaine de pages, il constitue une introduction remarquable à la topologie différentielle pour tout lecteur qui a maîtrisé les quatre premiers chapitres du présent volume.

Pour indiquer à quel point le sujet de topologie différentielle est encore assez ouvert, mentionnons le problème « élémentaire » suivant : si U, V sont deux parties ouvertes dans $\mathbb{R}^N$ qui sont homéomorphes, sont-elles aussi C^∞-difféomorphes ? Ceci est effectivement le cas si $N \leqslant 3$ mais le problème semble être non résolu pour $N \geqslant 4$; même pour $N \leqslant 3$, la démonstration requise n'est pas facile (cf. article [10] de Munkres dans son ouvrage [27]). Il y a un résultat de Stallings (*Proceedings of the Cambridge Philosophical Society* (1962)) selon lequel si U est homéomorphe à $\mathbb{R}^N$ et $N \geqslant 5$ alors U est difféomorphe à $\mathbb{R}^N$. Les travaux récents de Donaldson, Freedman et d'autres semblent indiquer que la dimension $N = 4$ peut être une dimension exceptionnelle pour le problème posé ; néanmoins, ce fait ne semble pas être établi. (Un article récent de M. H. Freedman et S. DeMechilis (*Journal of differential geometry*, vol. 35, 1992) démontre qu'il existe une infinité d'ensembles ouverts dans $\mathbb{R}^4$ qui sont homéomorphes à $\mathbb{R}^4$ mais qui ne sont pas difféomorphes entre eux.) Je remercie M. Munkres pour une correspondance très instructive au sujet du problème mentionné. En passant, je recommande vivement la lecture de son texte élémentaire (Munkres [28]) qui traite les questions d'analyse dans les variétés, à un niveau tout à fait comparable au présent volume.

CHAPITRE 4

Fonctions implicites

Il y a trois propositions centrales dans ce chapitre : théorème d'inversion locale, théorème des fonctions implicites, théorème du rang constant ; les deux dernières sont déduites de la première. Ces propositions forment l'outillage principal pour toutes les discussions concernant l'existence des applications de différents types nécessaires en analyse différentielle. Les sous-variétés différentielles de $\mathbb{R}^N$ sont introduites précisément ; elles joueront un rôle essentiel dans la suite, surtout celles de dimension 1 (les courbes) et de dimension 2 (les surfaces). On présente une discussion de la notion d'indépendance fonctionnelle et de la méthode de Lagrange pour les extremums liés.

On conseille au débutant de bien lire la section 4.1 et de s'assurer qu'il est capable de dériver les fonctions implicites en faisant les exercices du paragraphe 4.1.4. Il doit ensuite essayer de résoudre les premiers exercices des paragraphes 4.2.10 et 4.3.3 pour tester sa compréhension des énoncés des théorèmes d'inversions locale et des fonctions implicites.

4.1 PRÉLIMINAIRES

Dans cette section nous expliquons la position du problème traité dans les deux théorèmes principaux de ce chapitre : le ***théorème d'inversion locale***, le ***théorème des fonctions implicites***. Les formulations exactes de ces deux théorèmes et leurs démonstrations sont données dans les sections suivantes. Notre objectif immédiat est d'expliquer le contenu intuitif de ces théorèmes et d'indiquer leur plausibilité.

4.1.1 Position du problème

En termes simples, le problème est celui de la possibilité de la résolution d'un système d'équations

$$\begin{gathered} F_1(x_1, \ldots, x_N ; u_1, \ldots, u_M) = 0 \\ \qquad\qquad (4.1) \\ F_M(x_1, \ldots, x_N ; u_1, \ldots, u_M) = 0 \end{gathered}$$

où $F_1, \ldots, F_M$ sont M fonctions de $N+M$ variables réelles, notées $x_1, \ldots, x_N$, $u_1, \ldots, u_M$ et la solution désirée est de la forme

$$u_i = f_i(x_1, \ldots, x_N), \quad 1 \leqslant i \leqslant M .$$

Pour faciliter l'écriture, posons

$$x = (x_1, \ldots, x_N), \quad u = (u_1, \ldots, u_M), \quad F = (F_1, \ldots, F_M) ;$$

il s'agit donc de résoudre l'équation

$$F(x; u) = 0 \tag{4.2}$$

dans la forme $u = f(x)$, $f = (f_1, \ldots, f_M)$ de sorte que

$$F\big(x; f(x)\big) = 0 . \tag{4.3}$$

Si la fonction f existe, on dit qu'elle est définie ***implicitement*** par le système d'équation (4.1) (ou (4.2)). Un cas particulier important est celui de $M = N$ et de $F(x; u)$ de la forme $\varphi(x) - u$; autrement dit, il s'agit dans ce cas de résoudre l'équation

$$\varphi(x) = u \tag{4.4}$$

dans la forme $x = \psi(u)$ de sorte que

$$\varphi\big(\psi(u)\big) = u . \tag{4.5}$$

Ecrit de façon plus explicite, si $\varphi = (\varphi_1, \ldots, \varphi_N)$, on veut résoudre le système d'équations

$$\varphi_i(x_1, \ldots, x_N) = u_i , \quad 1 \leqslant i \leqslant N \tag{4.6}$$

dans la forme

$$x_i = \psi_i(u_1, \ldots, u_N) , \quad 1 \leqslant i \leqslant N .$$

La fonction $\psi = (\psi_1, \ldots, \psi_N)$ ainsi obtenue s'appelle la ***fonction inverse*** ou ***réciproque*** de φ ; on a alors

$$\varphi\big(\psi(u)\big) = u , \quad \psi\big(\varphi(x)\big) = x ;$$

on écrit aussi $\psi = \varphi^{-1}$ ou $\varphi = \psi^{-1}$.

4.1.2 Solution plausible du problème

Supposons que les F_i dans (4.1) soient des fonctions affines de $u_1, \ldots, u_M$, c'est-à-dire de la forme

$$F_i(x_1, \ldots, x_N ; u_1, \ldots, u_M) = \sum_{j=1}^{M} a_{ij}(x) u_j - c_i , \quad 1 \leqslant i \leqslant M$$

de sorte que

$$F(x; u) = A(x)u - c , \quad A(x) = \big[a_{ij}(x)\big] , \quad c = (c_1, \ldots, c_M) .$$

Le problème devient alors celui de résoudre l'équation linéaire

$$A(x)u = c\,.$$

Il est bien connu que cette équation possède une unique solution si $\det A(x) \neq 0$ et alors

$$u = [A(x)]^{-1}c\,.$$

Notons que dans ce cas, la matrice $A(x)$ est donnée par

$$A(x) = \left[\frac{\partial F_i}{\partial u_j}(x;u)\right]$$

et

$$\det A(x) = \frac{\partial(F_1, \ldots, F_M)}{\partial(u_1, \ldots, u_M)}(x;u)\,.$$

Nous démontrons dans la suite que, dans *le cas général*, si l'équation (4.1) possède au moins une solution $x = a$, $u = b$ et si la matrice fonctionnelle

$$\left[\frac{\partial F_i}{\partial u_j}(a;b)\right]$$

est inversible, c'est-à-dire le déterminant fonctionnel

$$\frac{\partial(F_1, \ldots, F_M)}{\partial(u_1, \ldots, u_M)}(a;b) \tag{4.7}$$

est non nul, alors (4.1) possède une unique solution $f(x)$ pour toutes les valeurs de x dans un voisinage ouvert de a et qu'il satisfait $f(a) = b$; pour démontrer ce type de proposition, nous supposerons que les fonctions F_i sont de classe C^1. De même, pour la solution de (4.4), nous supposerons que φ est de classe C^1 et que pour un point a fixé, $\varphi'(a)$ est inversible ; alors nous démontrerons l'existence d'une unique $\psi(u)$ définie dans un voisinage ouvert de $b = \varphi(a)$ telle que (4.5) est satisfait et telle que $\psi(b) = a$.

4.1.3 Calcul de la dérivée des fonctions implicites

Supposons que F dans (4.2) est de classe C^1 dans un voisinage ouvert de $(a;b) \in \mathbb{R}^{N+M}$ et que la fonction implicite f, définie dans un voisinage ouvert de $a \in \mathbb{R}^N$ via l'équation (4.3), est de classe C^1 dans un voisinage ouvert de a. Notre objectif est de calculer $f'(x)$ pour x dans un voisinage ouvert de a en supposant que le déterminant fonctionnel (4.7) est non nul ; cette dernière hypothèse entraîne alors que le même déterminant fonctionnel est non nul lorsque l'on remplace a par x dans un voisinage ouvert de a, et b par u dans un voisinage ouvert de b.

Posons $\Phi(x) = \big(x; f(x)\big)$; alors

$$F\big(x; f(x)\big) = F\big(\Phi(x)\big) = 0$$

dans un voisinage ouvert de a. Ainsi

$$\big(F \circ \Phi\big)'(x) = F'\big(\Phi(x)\big)\Phi'(x) = 0\,. \tag{4.8}$$

Posons

$$F'_x(x;u) = \left[\frac{\partial F_i}{\partial x_j}(x;u)\right], \quad F'_u(x;u) = \left[\frac{\partial F_i}{\partial u_j}(x;u)\right] \tag{4.9}$$

où F'_x est une matrice de type (M,N) et F'_u est une matrice de type (M,M) ; il est clair que

$$F'(x;u) = \big[F'_x(x;u),\ F'_u(x;u)\big]$$

une matrice de type $(M, N+M)$ et que

$$\Phi'(x) = \begin{bmatrix} I_N \\ f'(x) \end{bmatrix}$$

une matrice de type $(N+M, N)$ où $f'(x)$ est une matrice de type (M,N), I_N étant la matrice (carrée) identité de type (N,N). De (4.8) on aura alors

$$F'_x\big(x; f(x)\big) + F'_u\big(x; f(x)\big) f'(x) = 0\,, \tag{4.10}$$

d'où

$$f'(x) = -\big\{F'_u\big(x; f(x)\big)\big\}^{-1} F'_x\big(x; f(x)\big)\,. \tag{4.11}$$

Lorsque $M = N = 1$ les équations (4.10) et (4.11) sont familières et immédiates de la relation

$$F\big(x; f(x)\big) = 0\,;$$

le calcul précédent montre que ces équations sont valables dans la situation générale. Les formules (4.10) et (4.11) permettent le calcul de $f'(x)$ sans obtenir explicitement $f(x)$.

En ce qui concerne la dérivée de $\psi = \varphi^{-1}$, on sait déjà qu'elle se déduit de

$$\varphi\big(\psi(u)\big) = u$$

ce qui donne

$$\varphi'\big(\psi(u)\big)\psi'(u) = I_M\,, \quad \psi'(u) = \big\{\varphi'\big(\psi(u)\big)\big\}^{-1} \tag{4.12}$$

en supposant que φ' est inversible.

Dans les exercices qui suivent on aura la possibilité de pratiquer les calculs indiqués ci-dessus. L'existence des fonctions implicites et inverses dans ces exercices est admise en hypothèse ; l'objectif de ces exercices est de calculer leurs dérivées ; leur existence, dans des voisinages ouverts appropriés, est garantie grâce aux théorèmes des fonctions implicites et des fonctions inverses qui seront donnés dans la suite.

4.1.4 Exercices

[1] Considérer y comme une fonction de x définie par $x^2 + y^2 = c^2$ (c constante > 0) avec $y(0) = -c$; calculer $y'(x)$ par le procédé de la dérivation implicite du paragraphe 4.1.3 et comparer votre résultat avec celui obtenu du fait que $y(x) = -\sqrt{c^2 - x^2}$, $-c < x < c$.

[2] Si $z = f(x, y)$ est définie comme une solution de l'équation

$$e^z(x^2 + y^2 + z^2) - (1 + z^2)^{1/2} + y = 0$$

avec $f(1, 0) = 0$, écrire le développement de Taylor d'ordre 1 de f autour de $(1, 0)$.

[3] Si $y = f(x)$ est définie comme une solution de l'équation $y^3 = e^{x+y}$ avec $f(-1) = 1$ calculer $f'(-1)$.

[4] Si $y = f(x)$, $z = g(x)$ sont définies comme solutions du système

$$x^2 + \frac{y^2}{2} + z^3 - z^2 = \frac{3}{2}$$
$$x^3 + y^3 - 3y + z = -3$$

avec $f(-1) = 1$, $g(-1) = 0$ calculer $f'(-1)$, $g'(-1)$.

[5] Si $z = f(x, y, \omega)$ est définie comme une solution de l'équation

$$\omega e^{x-y} + x e^{y-z} + y e^{z-\omega} + z e^{\omega - x} = 4$$

avec $f(1, 1, 1) = 1$ calculer $\nabla f(1, 1, 1)$.

[6] Si l'équation $F(x, y, z) = 0$ est telle que l'on peut résoudre (différentiablement) n'importe quelle variable en termes des deux autres, démontrer que

$$\frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial y} \cdot \frac{\partial y}{\partial z} = -1$$

en interprétant cette formule correctement.

[7] Si $u = f(x, y, z)$, $v = g(x, y, z)$ sont des solutions du système d'équations

$$xy^2 + xzu + yv^2 = 3$$
$$u^3yz + 2xv - u^2v^2 = 2$$

telles que $f(1, 1, 1) = g(1, 1, 1) = 1$, calculer $\nabla f(1, 1, 1)$ et $\nabla g(1, 1, 1)$.

4.2 FONCTIONS INVERSES

L'objectif de cette section est d'établir un théorème garantissant l'existence ***locale*** d'une fonction inverse. Avant d'énoncer le théorème, établissons un peu de terminologie.

4.2.1 Terminologie

Soit $\varphi : A \to B$, $A \subset \mathbb{R}^N$, $B \subset \mathbb{R}^M$; nous dirons que φ est une ***injection locale autour du point*** $a \in A$ s'il existe un voisinage ouvert V de a tel que la restriction de φ à $A \cap V$ est injective ; nous dirons que φ est un ***homéomorphisme local autour du point*** $a \in A$ s'il existe un voisinage ouvert V de a et un voisinage ouvert W de $b = \varphi(a)$ tels que la restriction de φ à $A \cap V$ est un homéomorphisme entre $A \cap V$ et $B \cap W$; si les voisinages V, W peuvent être choisis de telle manière que la restriction de φ à $A \cap V$ est un C^k-difféomorphisme entre $A \cap V$ et $B \cap W$, alors nous dirons que φ est un ***C^k-difféomorphisme local autour du point*** $a \in A$. Il est clair que, dans ces définitions, un des voisinages ouverts V (resp. W) peut être pris comme une boule ouverte centrée en a (resp. centrée en b). Nous dirons que $\varphi : A \to B$ est une ***injection locale*** ou un ***homéomorphisme local*** ou un ***C^k-difféomorphisme local*** si φ possède la même propriété autour de chaque point $a \in A$; notons que si φ est un C^k-difféomorphisme local alors φ est aussi un homéomorphisme local ce qui entraîne l'injectivité locale de φ.

4.2.2 Proposition (*théorème des fonctions inverses ou théorème d'inversion locale*)

Soit $\varphi : \Omega \to \mathbb{R}^N$ une fonction de classe C^k, $1 \leqslant k \leqslant \infty$, définie dans un sous-ensemble ouvert Ω de $\mathbb{R}^N$. Si pour un point $a \in \Omega$, $\varphi'(a)$ est inversible alors φ est un C^k-difféomorphisme local autour de a.

Nous démontrerons cette proposition comme une conséquence de deux lemmes dont chacun établit une partie de la proposition énoncée.

4.2.3 Lemme (*injectivité*)

Soit $\varphi : \Omega \to \mathbb{R}^N$ une fonction de classe C^1 définie dans un sous-ensemble ouvert Ω de $\mathbb{R}^N$. Si pour un point $a \in \Omega$, $\varphi'(a)$ est inversible alors φ est une injection locale autour de a.

DÉMONSTRATION

Soit $M = [m_{ij}]$ la matrice canonique pour $\varphi'(a)$, c'est-à-dire $m_{ij} = \mathrm{D}_j\varphi_i(a)$ où $\varphi = (\varphi_1, \ldots, \varphi_N)$.

Si $\varphi'(a)$ est inversible, on aura $\det M \neq 0$; il existe alors un $\varepsilon > 0$ tel que toute matrice $P = [p_{ij}]$ de type (N, N) ayant la propriété

$$|p_{ij} - m_{ij}| < \varepsilon\,, \quad 1 \leqslant i, j \leqslant N$$

sera aussi inversible (c'est-à-dire $\det P \neq 0$) ; c'est une simple conséquence du fait que le déterminant est une fonction continue de ses N^2 termes. Puisque φ est de classe C^1, on peut choisir un $r > 0$ tel que $B = B(a; r) \subset \Omega$ et

$$\left|\mathrm{D}_j\varphi_i(x) - \mathrm{D}_j\varphi_i(a)\right| < \varepsilon\,, \quad 1 \leqslant i, j \leqslant N\,,\ x \in B\,.$$

Alors $\varphi'(x)$ sera inversible si $x \in B$. Si x, y sont dans B, il existe, pour chaque i, $1 \leqslant i \leqslant N$, un $\xi_i \in B$ (dans le segment ouvert $]\,x\,,\,y\,[\,$) tel que

$$\varphi_i(x) - \varphi_i(y) = (x - y) \cdot \nabla\varphi_i(\xi_i) = \left[\mathrm{D}_1\varphi_i(\xi_i), \ldots, \mathrm{D}_N\varphi_i(\xi_i)\right](x - y)$$

de sorte que

$$\varphi(x) - \varphi(y) = P(x - y)$$

où $P = [p_{ij}]$, $p_{ij} = \mathrm{D}_j\varphi_i(\xi_i)$; puisque

$$|p_{ij} - m_{ij}| < \varepsilon\,, \quad 1 \leqslant i, j \leqslant N\,,$$

on aura que P est inversible ce qui entraîne que si $x \neq y$ sont deux points de B alors

$$\varphi(x) - \varphi(y) = P(x - y) \neq 0\,;$$

ceci veut dire que φ est injective dans B. ♦

4.2.4 Lemme (*application ouverte*)

Soit $\varphi : \Omega \to \mathbb{R}^N$ une fonction injective de classe C^1 définie dans une partie ouverte Ω de $\mathbb{R}^N$; supposons que $\varphi'(x)$ est inversible pour tout $x \in \Omega$. Alors φ est une application ouverte.

DÉMONSTRATION

Il faut voir que $\varphi(A)$ est ouvert si $A \subset \Omega$ est ouvert ; il suffit en fait de démontrer que $\Omega' = \varphi(\Omega)$ est ouvert. Soit $a \in \Omega$ et posons $b = \varphi(a)$; il faut établir l'existence d'un $r > 0$ tel que $B(b; r) \subset \Omega'$.

Soit B une boule ouverte de centre a telle que $\overline{B} \subset \Omega$; posons $S = \mathrm{Fr}\,B$. Puisque S est compact et φ est continue, $\varphi(S)$ est compact ; φ étant injective on a que $b \notin \varphi(S)$ car déjà $b = \varphi(a)$ et $a \notin S$. Choisissons $r > 0$ tel que

$$\inf_{x \in S} \left\|b - \varphi(x)\right\| = d\big(b; \varphi(S)\big) > 2r > 0\,;$$

vérifions que $B(b; r) \subset \varphi(B) \subset \Omega'$.

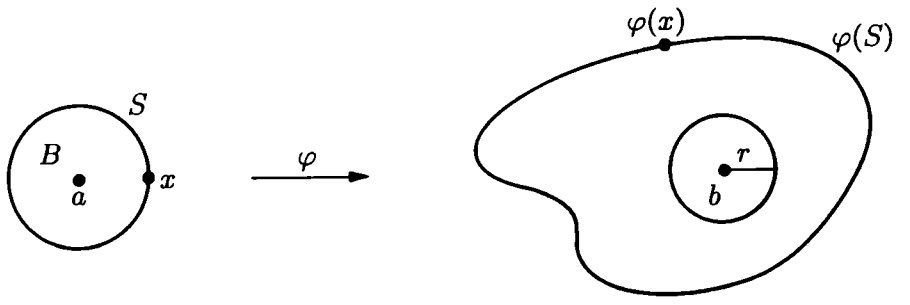


Fig. 4.1

En effet, soit c tel que $\|b-c\| < r$; posons

$$F(x) = \|\varphi(x) - c\|^2 , \quad x \in \Omega .$$

Il est clair que F est de classe C^1 et que F atteint son infimum dans $\overline{B}$ en un point ξ de $\overline{B}$, c'est-à-dire

$$F(\xi) = \inf_{x \in \overline{B}} F(x) , \quad \overline{B} = B \cup S .$$

Démontrons que $\xi \in B$ et $F(\xi) = 0$; en effet, on a

$$F(a) = \|\varphi(a) - c\|^2 = \|b - c\|^2 < r^2$$

d'où $F(\xi) < r^2$; d'autre part, si $x \in S$, on a

$$\|\varphi(x) - b\| > 2r$$

d'où

$$\|\varphi(x) - c\| \geqslant \|\varphi(x) - b\| - \|b - c\| > r$$

ce qui montre que $F(x) > r^2$ si $x \in S$. On en conclut que $\xi \notin S$, ce qui veut dire que $\xi \in B$; ainsi ξ est un minimum relatif pour la fonction F, ce qui entraîne que

$$F'(\xi) = \nabla F(\xi) = 0 .$$

Or, $F(x) = f\big(\varphi(x) - c\big)$ avec $f(x) = \|x\|^2$ et $f'(x) = 2 \cdot {}^t x$ d'où

$$F'(x) = f'\big(\varphi(x) - c\big)\varphi'(x) = 2 \cdot {}^t\big(\varphi(x) - c\big)\varphi'(x) ;$$

ainsi, si $F'(\xi) = 0$ on aura $\varphi(\xi) - c = 0$ car $\varphi'(\xi)$ est inversible. Au total, si $c \in B(b;r)$ on a un $\xi \in B$ tel que $\varphi(\xi) = c$, ce qui démontre que

$$B(b;r) \subset \varphi(B) . \qquad \blacklozenge$$

4.2.5 Démonstration de la proposition 4.2.2

Selon le lemme 4.2.3, il existe une boule ouverte V de centre a contenue dans Ω telle que φ est injective dans V ; puisque φ est continûment dérivable et

$\det \varphi'(a) \neq 0$, nous pouvons choisir V telle que $\det \varphi'(x) \neq 0$ pour tout $x \in V$. Le lemme 4.2.4 permet maintenant de conclure que $W = \varphi(V)$ est un voisinage ouvert de $b = \varphi(a)$; de plus, puisque φ transforme toute partie ouverte contenue dans V en une partie ouverte contenue dans W et que φ est une bijection continue de V sur W, on a que $\varphi : V \to W$ est un homéomorphisme tel que $\varphi'(x)$ est inversible pour tout $x \in V$. Selon la proposition 3.9.3, φ est un C^k-difféomorphisme entre V et W. ♦

4.2.6 Proposition

Soit $\varphi : \Omega \to \mathbb{R}^N$ une fonction continue définie dans une partie ouverte Ω de $\mathbb{R}^N$. Si φ est un homéomorphisme local alors φ est une application ouverte. En particulier, si φ est de classe C^1 et $\det \varphi'(x) \neq 0$ pour tout $x \in \Omega$ alors φ est une application ouverte.

DÉMONSTRATION

La première affirmation est une conséquence immédiate de la définition d'un homéomorphisme local. La seconde en est une conséquence compte tenu de la proposition 4.2.2. ♦

4.2.7 Remarques

Mentionnons ici un théorème important dû à Brouwer (1912) que l'on nomme théorème de l'***invariance du domaine*** : *soit $\varphi : \Omega \to \mathbb{R}^N$ une fonction continue et injective définie dans une partie ouverte Ω de $\mathbb{R}^N$; alors φ est une application ouverte ; ainsi, $\Omega' = \varphi(\Omega)$ est ouverte et $\varphi : \Omega \to \Omega'$ est un homéomorphisme.*

On a déjà vu que cette proposition est en défaut si Ω n'est pas ouvert (Ex. 7, § 1.6.4). L'appellation «invariance du domaine» provient du fait que, sous l'hypothèse imposée sur φ, l'image $\varphi(A)$ d'un domaine $A \subset \Omega$ est de nouveau un domaine dans $\mathbb{R}^N$; en effet, le domaine A, étant une partie (non vide) ouverte et connexe, est transformé par φ à une partie (non vide) $\varphi(A)$ qui est ouverte (selon la proposition citée de Brouwer) et connexe (car φ est continue) selon la proposition élémentaire 1.8.2(ii). La démonstration de ce théorème de Brouwer demande des considérations que nous ne développerons pas ici ; une preuve assez élémentaire est donnée dans le livre de Hurewicz et Wallman [22, p. 95] ; voir aussi l'excellent petit livre de Newman [31, p. 122] qui expose plusieurs aspects avancés de la topologie de $\mathbb{R}^N$ avec des moyens relativement élémentaires. Remarquons que certains auteurs énoncent le théorème de l'invariance du domaine sous l'hypothèse préliminaire plus forte sur φ ; ils partent de l'hypothèse que φ est déjà un homéomorphisme entre Ω et $\varphi(\Omega)$; même avec cette hypothèse forte, il n'est pas du tout facile d'établir que $\varphi(\Omega)$ est ouvert si Ω est ouvert dans $\mathbb{R}^N$. Nous pensons que la formulation du théorème de l'invariance du domaine, donnée ici, est plus commode pour les applications en analyse. Néanmoins, les deux formulations du théorème sont équivalentes ; on

peut démontrer une des versions à partir de l'autre assez facilement. En effet, la version où l'on suppose que φ est déjà un homéomorphisme est une conséquence évidente de la formulation du théorème donnée ici. Inversement, si l'on admet la version où φ est supposée d'emblée être un homéomorphisme, alors on en déduit immédiatement qu'une application continue et injective $\varphi : \Omega \to \mathbb{R}^N$ (Ω étant une partie ouverte de $\mathbb{R}^N$) est un homéomorphisme local (car $\varphi|_{\overline{B}}$ est alors un homéomorphisme de $\overline{B}$ sur $\varphi(\overline{B})$ si B est une boule ouverte telle que $\overline{B} \subset \Omega$), d'où φ est une application ouverte (prop. 4.2.6).

Une conséquence importante du théorème de l'invariance du domaine est la proposition suivante : *si $M < N$ alors on ne peut pas avoir une application continue et injective $\varphi : \Omega \to \mathbb{R}^M$ définie dans une partie ouverte non vide $\Omega \subset \mathbb{R}^N$*. En effet, si $\psi : \mathbb{R}^M \to \mathbb{R}^N$ est l'injection suivante :

$$\psi(x_1, \ldots, x_M) = (x_1, \ldots, x_M, 0, \ldots, 0),$$

alors $\psi \circ \varphi$ serait une application continue et injective de Ω dans $\mathbb{R}^N$ si $\varphi : \Omega \to \mathbb{R}^M$ était continue et injective ; or, l'image Ω' de Ω sous $\psi \circ \varphi$ est contenue dans un sous-espace linéaire de $\mathbb{R}^N$ de dimension $M < N$ de sorte que Ω' ne peut pas être ouverte (en fait, $\mathring{\Omega}' = \phi$), ce qui contredit le théorème de l'invariance du domaine.

Une autre conséquence du théorème de l'invariance du domaine est la proposition suivante : *si $M > N$ et $\varphi : \Omega \to \mathbb{R}^M$ est une application continue et injective définie dans une partie ouverte Ω de $\mathbb{R}^N$, alors l'ensemble $\varphi(\Omega)$ dans $\mathbb{R}^M$ est d'intérieur vide*. La preuve serait assez immédiate si l'on supposait que φ est un homéomorphisme de Ω sur $\varphi(\Omega)$ car, si W est une partie ouverte non vide contenue dans $\varphi(\Omega)$ et $V = \varphi^{-1}(W)$, alors $\varphi^{-1}|_W$ donnerait un homéomorphisme entre W et V, $W \subset \mathbb{R}^M$, $V \subset \mathbb{R}^N$, $N < M$, contredisant la proposition précédente. La preuve, sous l'hypothèse que φ est seulement continue et injective, est un petit peu plus subtile ; notons, en passant, que cette hypothèse sur φ n'entraîne pas que φ est un homéomorphisme entre Ω et $\varphi(\Omega)$ (car φ^{-1} n'est pas nécessairement continue, Ex. 1, § 4.2.10). Néanmoins, $\varphi|_{\overline{B}}$ est un homéomorphisme entre $\overline{B}$ et $\varphi(\overline{B})$ si B est une boule dans Ω telle que $\overline{B} \subset \Omega$ (car $\overline{B}$ est compact) ; selon la version homéomorphique déjà établie, on aura que $\varphi(\overline{B})$ possède un intérieur vide. Si W est une partie ouverte non vide contenue dans $\varphi(\Omega)$, alors $V = \varphi^{-1}(W)$ est une partie ouverte contenue dans Ω ; en écrivant V comme une réunion dénombrable des boules fermées $\overline{B}_n$, on aura

$$V = \bigcup_{n=1}^{\infty} \overline{B}_n, \quad W = \varphi(V) = \bigcup_{n=1}^{\infty} \varphi(\overline{B}_n)$$

ce qui donne la partie ouverte $W \subset \mathbb{R}^M$ comme une réunion dénombrable des ensembles fermés $\varphi(\overline{B}_n)$, chacun d'intérieur vide ; un théorème de Baire (que nous énonçons et démontrons dans la forme nécessaire ici dans l'exercice 13 du paragraphe 4.2.10) ne permet pas de telle représentation pour une partie ouverte non vide W de $\mathbb{R}^M$.

Pour compléter notre discussion mentionnons qu'il existe bien des fonctions continues, *non injectives*, $\varphi : \Omega \to \mathbb{R}^M$, Ω ouvert dans $\mathbb{R}^N$, $M > N$, telles que $\varphi(\Omega)$ contienne des ouverts non vides. Peano a donné (en 1890) une application continue de $[0, 1]$ remplissant un carré de $\mathbb{R}^2$ et aujourd'hui on sait qu'étant donné une boule fermée quelconque B dans $\mathbb{R}^M$ ($M \geqslant 1$ arbitraire), il existe une application continue $\varphi : I \to \mathbb{R}^M$, $I = [0, 1] \subset \mathbb{R}$, telle que $\varphi(I) = B$; voir le théorème précis dû à Hahn et Mazurkiewicz dans Newman [31, p. 89] ; on dit que φ *définit une courbe de Peano*. Il est intéressant de remarquer que la seule restriction d'injectivité sur φ élimine ce phénomène contre-intuitif d'avoir une « courbe » de « dimension » quelconque. En fait, on peut donner une bonne définition de la notion de *dimension topologique* d'un ensemble $A \subset \mathbb{R}^N$ et démontrer que si $\mathring{A} \neq \emptyset$ alors la dimension topologique de A est N (Hurewicz et Wallman [22] pour une excellente introduction à ce sujet). De notre discussion on voit immédiatement que si V est une partie ouverte de $\mathbb{R}^M$, W est une partie ouverte de $\mathbb{R}^N$, et si V et W sont homéomorphes alors on a $M = N$; c'est *l'invariance de la dimension* sous homéomorphisme. Notons que l'invariance de la dimension sous difféomorphisme est presque immédiate (Ex. 2, § 4.2.10), alors que l'invariance de la dimension sous homéomorphisme semble être beaucoup plus difficile à démontrer.

En admettant le théorème de l'invariance du domaine, on peut donner une démonstration très courte de la proposition 4.2.2 (théorème d'inversion locale). En effet, le lemme simple 4.2.3 dit que φ est localement injective au point a si $\varphi'(a)$ est inversible, d'où on conclut immédiatement, grâce au théorème de l'invariance du domaine, que φ est un homéomorphisme local au point a ; on en déduit, grâce à la proposition 3.9.3, que φ est un C^k-difféomorphisme local au point a si $\det \varphi'(a) \neq 0$.

4.2.8 Les images

Etant donnée une application $\varphi : X \to \mathbb{R}^N$, il est important de déterminer l'ensemble image $\varphi(X)$. Supposons que $X = \overline{\Omega}$ où Ω est une partie *ouverte bornée* dans $\mathbb{R}^N$; si $K = \operatorname{Fr} \Omega$, $\overline{\Omega} = \Omega \cup K$; supposons que $\varphi : \overline{\Omega} \to \mathbb{R}^N$ est *continue et* $\varphi|_\Omega$ *est injective*, alors

$$\varphi(\operatorname{Fr} \Omega) = \operatorname{Fr} \varphi(\Omega) . \tag{4.13}$$

Pour démontrer (4.13), posons $\Omega' = \varphi(\Omega)$, $K' = \operatorname{Fr} \Omega'$; selon le théorème de l'invariance du domaine, Ω' est ouvert.

Démontrons d'abord que $\varphi(K) \subset K'$. Soient $\xi \in K$, $\varphi(\xi) = \eta$, alors $\eta \notin \Omega'$. Sinon, $\eta \in \Omega'$ et il existe $a \in \Omega$ avec $\varphi(a) = \eta$; prenons $a_n \in \Omega$, $n = 1, 2, \ldots$ tels que $a_n \to \xi$ lorsque $n \to \infty$ et posons $b_n = \varphi(a_n) \in \Omega'$; puisque φ est continue, $b_n \to \varphi(\xi) = \eta$ lorsque $n \to \infty$; soit V un voisinage ouvert de a contenu dans Ω tel que $a_n \notin V$, $n \geqslant n_0$; alors $W = \varphi(V)$ est un voisinage ouvert de $\eta = \varphi(a)$ contenu dans Ω' ; or, $b_n \to \eta$ ($n \to \infty$) entraîne que $b_n \in W$ pour $n \geqslant n_1$ ce qui donne que $a_n \in V$ pour $n \geqslant n_1$ contredisant le

choix de V. Donc $\eta \notin \Omega'$; en fait, $\eta \in K'$ car η est la limite des $b_n \in \Omega'$; ainsi $\varphi(K) \subset K'$.

Démontrons maintenant que $K' \subset \varphi(K)$. Soit $\eta \in K'$; il existe $b_n \in \Omega'$, $n = 1, 2, \ldots$, tels que $b_n \to \eta$; soit $a_n \in \Omega$ avec $\varphi(a_n) = b_n$; si une sous-suite des a_n converge vers $\xi \in \overline{\Omega}$, $\xi \notin \Omega$ car sinon $\xi \in \Omega$ et $\eta = \varphi(\xi)$ sera dans Ω' ; comme $\overline{\Omega}$ est compact, il existe au moins une sous-suite des a_n qui converge vers un point $\xi \in K$ d'où $\eta = \varphi(\xi)$ donne que $\eta \in \varphi(K)$; ainsi, $K' \subset \varphi(K)$ et (4.13) est démontrée.

Notons que la relation (4.13) subsiste aussi si Ω n'est pas borné pourvu que nous imposons sur φ l'hypothèse supplémentaire suivante : $|\varphi(x)| \to \infty$ lorsque $|x| \to \infty$; la démonstration précédente est applicable avec à la fin un léger changement que nous laissons comme exercice.

L'intérêt pratique de (4.13) réside dans l'observation suivante. Supposons, par exemple, que $K = \operatorname{Fr} \Omega$ est une « courbe simple » formant la frontière de $\Omega \subset \mathbb{R}^2$; alors il suffit de déterminer la « courbe image » $K' = \varphi(K)$ pour dessiner $\Omega' = \varphi(\Omega)$ dont la frontière est K' ; notons que

$$\varphi(\overline{\Omega}) = \varphi(\Omega \cup K) = \varphi(\Omega) \cup \varphi(K) = \Omega' \cup K' .$$

Dans $\mathbb{R}^3$, on travaillera avec les « surfaces » $K = \operatorname{Fr} \Omega$, $K' = \operatorname{Fr} \Omega'$.

Rappelons que (4.13) est démontrée sous des hypothèses fortes sur φ (injectivité de φ en particulier) ; en général, la relation entre $\varphi(\operatorname{Fr} \Omega)$ et $\operatorname{Fr} \varphi(\Omega)$ est loin d'être simple.

4.2.9 Théorèmes d'inversion globale

La proposition 4.2.2 est un théorème qui permet de définir une fonction inverse localement ; même si on renforce les hypothèses de cette proposition en supposant que $\varphi'(x)$ est inversible pour tout $x \in \Omega$, on ne peut pas en déduire l'existence globale d'une fonction inverse, c'est-à-dire s'assurer que φ est injective partout dans Ω. Les exemples simples de ce fait sont donnés dans les exercices du paragraphe 4.2.10. Néanmoins, il est important de savoir sous quelles conditions simples sur φ et Ω nous pouvons affirmer que φ^{-1} existe. L'objectif de ce paragraphe est d'indiquer quelques résultats utiles, sans démonstration, de ce type. Malheureusement, malgré l'existence d'une recherche assez étendue autour du problème, il n'y a pas de textes élémentaires ou de monographies accessibles qui font un tour complet de ce problème important. Nous donnerons quelques références appropriées au moment opportum.

Un théorème très utile que nous donnons en premier lieu exige quelques notions topologiques nouvelles. Soit $\varphi : \Omega \to \mathbb{R}^M$, Ω ouvert dans $\mathbb{R}^N$; on dit que φ est une application ***propre*** si $\varphi^{-1}(K)$ est compact dès que $K \subset \mathbb{R}^M$ est compact. Une autre notion topologique dont nous aurons besoin est celle de ***simple connectivité*** ; nous développons cette notion plus longuement dans les chapitres suivants ; disons ici seulement qu'un sous-ensemble $A \subset \mathbb{R}^N$ s'appelle simplement connexe s'il est connexe par arcs et si tout chemin fermé dans A (c'est-à-dire une application continue $\gamma : [0, 1] \to A$ avec $\gamma(0) = \gamma(1)$) peut

être continûment déformé dans A à un chemin constant. Un ensemble ouvert et convexe est un exemple simple d'un ensemble simplement connexe. Le théorème en question s'énonce maintenant comme suit : *soient Ω, Ω' deux domaines de $\mathbb{R}^N$ et $\varphi : \Omega \to \Omega'$ une application continue localement inversible et propre ; si Ω' est simplement connexe alors φ est un homéomorphisme de Ω sur Ω'* (et, en particulier, une bijection entre Ω et Ω'). Même sans la condition de simple connexité sur Ω' on peut dire que, pour tout $y \in \Omega'$, le sous-ensemble $\varphi^{-1}(y)$ de Ω est d'une cardinalité finie et constante (indépendante de y) ; cette affirmation n'est pas difficile à établir ; le fait que $\varphi^{-1}(y)$ contient exactement un point pour tout $y \in \Omega'$ est plus subtile et exige une condition supplémentaire sur Ω'. Notons que sous les conditions du théorème énoncé, le domaine Ω sera également simplement connexe.

Une démonstration détaillée de ce théorème dans un cadre plus général (où Ω, Ω' sont deux espaces métriques appropriés) est donnée dans le livre de Ambrosetti et Prodi [1, p. 47]. Un cas particulier du théorème est le suivant : *soit $\varphi : \overline{E} \to \overline{F}$ une application continue où E, F sont deux domaines bornés, simplement connexes dans $\mathbb{R}^N$; si $\varphi|_E$ est une injection locale et*

$$\varphi(E) = F\,, \quad \varphi(\operatorname{Fr} E) \subset \operatorname{Fr} F \tag{4.14}$$

alors φ est un homéomorphisme de E sur F.

En effet, la condition (4.14), alliée aux autres hypothèses, fait que $\varphi|_E$ est une application propre de E dans F, ce qui permet de déduire ce résultat immédiatement du théorème précédent.

Un théorème d'un type différent (dû à Gale et Nikaido, 1965) est le suivant : *soit $f : A \to \mathbb{R}^N$ une application dérivable définie dans un pavé fermé $A \subset \mathbb{R}^N$ (borné ou non) ; supposons que, pour tout $x \in A$, la matrice canonique de $f'(x)$ est telle que tous ses mineurs principaux sont strictement positifs ; alors f est injective.* En particulier, si la matrice canonique de $f'(x)$ est symétrique et de type strictement positif alors f est injective ; ceci est un simple exercice (Ex. 20, sect. 2.7) valable sous l'hypothèse générale que A est une partie convexe ouverte quelconque. La généralisation du théorème de Gale et Nikaido au cas d'un sous-ensemble compact convexe A semble être inconnue ; pour une démonstration du théorème cité et pour d'autres résultats concernant l'inversion globale, on peut consulter la monographie de Parthasarathy [33]. Dans les exercices du paragraphe 4.2.10 on trouve différents exemples en relation avec les théorèmes donnés dans ce paragraphe.

Rappelons que si $[a_{ij}]$ est une matrice de type (N, N) , ses ***mineurs d'ordre*** p sont les nombres

$$\det \begin{bmatrix} a_{i_1 j_1} & \dots & a_{i_1 j_p} \\ \vdots & \ddots & \vdots \\ a_{i_p j_1} & \dots & a_{i_p j_p} \end{bmatrix}$$

où $1 \leqslant i_1 < \ldots < i_p \leqslant N$, $1 \leqslant j_1 < \ldots < j_p \leqslant N$, ses ***mineurs principaux*** (d'ordre p) correspondent aux mineurs où $j_1 = i_1, \ldots, j_p = k_p$; les ***mineurs***

principaux diagonaux correspondent aux mineurs pour lesquels $i_1 = j_1 = 1, \ldots, i_p = j_p = p$ que nous avons déjà définis dans le paragraphe 3.4.6.

Pour clore ce paragraphe, mentionnons un autre outil topologique puissant pour l'étude des problèmes globaux du type traité ici et pour d'autres questions importantes : la ***théorie du degré*** d'une application (due à Brouwer). On trouvera une introduction à cette théorie dans le livre de Milnor [26] déjà cité. Le célèbre ***théorème du point fixe*** (dû aussi à Brouwer) en est une conséquence assez directe ; pour l'information du lecteur nous le citons comme suit : *soit B une boule fermée dans $\mathbb{R}^N$; si $f : B \to B$ est une application continue alors f possède au moins un point fixe* (c'est-à-dire, il existe $x \in B$ tel que $f(x) = x$). Le théorème du point fixe de Brouwer est lié à plusieurs questions centrales de la topologie (traitées aussi dans Milnor [26]).

4.2.10 Exercices

[1] Soit $\varphi :]-1, \infty[\to \mathbb{R}^2$ une application définie comme suit : $\varphi(x) = (x, 0)$ si $x \geqslant 0$; $\varphi(x) = \big(-x, f(x)\big)$ si $-1 < x < 0$ où $f : [-1, 0] \to \mathbb{R}$ est une fonction continue telle que $f(0) = f(-1) = 0$, $f(x) > 0$ si $-1 < x < 0$ (par exemple $f(x) = -x(x+1)$). Vérifier que φ est une application continue et injective mais que φ^{-1} n'est pas continue. (Dessiner le chemin tracé par φ.) En écrivant $\Omega =]-1, \infty[$ on voit que $\varphi(\Omega)$ est d'intérieur vide dans $\mathbb{R}^2$; en prolongeant φ à $\overline{\Omega}$ par $\varphi(-1) = (1, 0)$ on aura $\varphi(\operatorname{Fr}\Omega) = (1, 0) \neq \operatorname{Fr}\varphi(\Omega)$.

[2] Soit $\varphi : \Omega \to \Omega'$ un difféomorphisme entre Ω ouvert dans $\mathbb{R}^N$ et Ω' ouvert dans $\mathbb{R}^M$; conclure directement que $M = N$. (Si $\psi = \varphi^{-1}$, $\psi \circ \varphi = \mathrm{id}$ d'où $\psi'\big(\varphi(x)\big)\varphi'(x) = I_N$, I_N matrice identité (N, N) d'où $\varphi'(x)$ inversible pour tout x ce qui n'est possible que pour $M = N$.)

[3] Vérifier que $u = x + x^2 + y$, $v = x^2 + y^2$ définit une application bijective $(x, y) \mapsto (u, v)$ dans un voisinage ouvert de $(x, y) = (5, 8)$.

[4] Soit $\varphi(r, \theta) = (r\cos\theta\,,\, r\sin\theta)$, $r > 0$, $\theta \in \mathbb{R}$. Vérifier que φ est une application de classe C^∞ qui est surjective de $]0, \infty[\times \mathbb{R}$ sur $\mathbb{R}^2 \setminus (0, 0)$; aussi $\varphi'(r, \theta)$ est inversible pour tout $(r, \theta) \in]0, \infty[\times \mathbb{R}$ mais φ n'est pas injective (cf. sect. 3.5, coordonnées polaires dans $\mathbb{R}^2$).

[5] Posons $\Omega =]0, \infty[\times]-\pi, \pi[$, $\Omega' = \mathbb{R}^2 \setminus \mathbb{R}_-$, $\mathbb{R}_- = \{(x, 0) : x \leqslant 0\}$; φ, défini dans l'exercice 4, donne une application de $\overline{\Omega}$ sur $\mathbb{R}^2$ telle que $\varphi(\Omega) = \Omega'$ mais $\varphi(\operatorname{Fr}\Omega) = \mathbb{R}_- \neq \operatorname{Fr}\mathbb{R}^2 = \emptyset$.

[6] Posons $\Omega =]0, 2[\times]0, 1[$, $\Omega' =]0, 1[\times]0, 1[$ et $\varphi(x, y) = (x, y)$ si $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ et $\varphi(x, y) = \big(\alpha(1 - x) + 1\,,\, y\big)$ si $1 < x \leqslant 2$, $0 \leqslant y \leqslant 1$ où $0 < \alpha < 1$. Vérifier que φ est une application continue de $\overline{\Omega}$ sur $\overline{\Omega'}$ mais que $\varphi(\operatorname{Fr}\Omega) \neq \operatorname{Fr}\varphi(\Omega)$; aussi, $(1, y)$ est un point intérieur de $\overline{\Omega}$ dont l'image par φ est un point de $\operatorname{Fr}(\overline{\Omega'}) = \operatorname{Fr}(\Omega')$. (L'application φ plie le rectangle fermé $\overline{\Omega}$ suivant l'axe $x = 1$ en faisant une compression appropriée pour rester dans $\overline{\Omega'}$.)

7 Soit $\varphi(x,y) = (x^2 - y^2\,,\, 2xy)$, $(x,y) \in \mathbb{R}^2$. Vérifier que $\varphi'(x,y)$ est inversible pour tout $(x,y) \neq (0,0)$ mais que φ en tant qu'application de $\mathbb{R}^2 \setminus (0,0)$ dans lui-même est surjective sans être injective ; de plus, φ n'est pas localement injective autour de $(0,0)$. Si $D = \{(x,y) : x > 0\}$ et $D' = \mathbb{R}^2 \setminus \mathbb{R}_-$, $\mathbb{R}_- = \{(x,0) : x \leqslant 0\}$, alors $\varphi : D \to D'$ est un homéomorphisme ; déterminer φ^{-1} explicitement. (L'application φ devient plus maniable dans son écriture en utilisant les nombres complexes : $\varphi(z) = z^2$ où $z = x + iy$.)

8 Vérifier directement que l'application $\varphi : D \to D'$ de l'exercice 7 est propre ; comme φ est localement inversible on déduit de nouveau que φ est un homéomorphisme grâce au théorème donné dans le paragraphe 4.2.9 en sachant que D' est simplement connexe.

9 Soit $\varphi(x,y) = (x^3/3 + x - y\,,\, -x + y)$; vérifier directement que $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ est une bijection (et un homéomorphisme) mais que $\varphi'(x,y)$ n'est pas inversible si $x^2 = 0$.

10 Soient $\varphi(x,y) = (\mathrm{e}^{2x} - y^2 + 3\,,\, 4y\mathrm{e}^{2x} - y^3)$, $(x,y) \in \mathbb{R}^2$, $A = [\,-2\,,\,2\,] \times [\,-2\,,\,2\,]$. Vérifier que les principaux mineurs diagonaux de $\varphi'(x,y)$ sont strictement positifs pour $(x,y) \in \mathbb{R}^2$ sans que $\varphi : A \to \mathbb{R}^2$ soit injective car $\varphi(0,2) = \varphi(0,-2) = (0,0)$; ceci ne contredit pas le théorème de Gale-Nikaido car un des mineurs principaux n'est pas strictement positif dans A. (Cet exemple est dû à Gale-Nikaido ; les exercices 9 et 10 sont tirés de Parthasarathy [33].)

11 Soit U une partie ouverte bornée de $\mathbb{R}^N$; si $\varphi = \overline{U} \to \mathbb{R}^M$ est une application continue telle que $\varphi|_U$ est injective et $\varphi(\mathrm{Fr}\,U) \cap \varphi(U) = \emptyset$, démontrer que $\varphi|_U : U \to \varphi(U)$ est un homéomorphisme. (Si $y_n \in \varphi(U)$, $y_n = \varphi(x_n)$, $x_n \in U$, $y_n \to y \in \varphi(U)$, $y = \varphi(x)$, $x \in U$, on montre que toute sous-suite convergente $\{x_{n_j}\}$ converge vers une même limite $x \in U$.)

12 Soit $\varphi : \mathbb{R}^N \to \mathbb{R}^N$ un homéomorphisme local ; si $|\varphi(x)| \to \infty$ lorsque $|x| \to \infty$ démontrer que φ est un homéomorphisme. (Vérifier d'abord que φ est une application propre et utiliser ensuite le théorème cité dans le paragraphe 4.2.9 ; le résultat de cet exercice s'appelle souvent le théorème de Hadamard ; plusieurs généralisations, dues à divers auteurs, sont données dans Ambrosetti et Prodi [1], Parthasarathy [33].)

13 Soit W une partie ouverte non vide de $\mathbb{R}^M$. Si $W = \bigcup_{n=1}^{\infty} A_n$ où les A_n sont fermés alors il existe un n tel que $\mathring{A}_n \neq \emptyset$. (Sinon, on aura $W = \bigcup_n A_n$, A_n fermé, $\mathring{A}_n = \emptyset$; il existe alors une boule fermée B_1 de rayon < 1 tel que $B_1 \subset W \cap A_1^c$; par récurrence, on obtient une suite de boule fermée $B_1 \supset B_2 \supset \dots$ telle que le rayon de $B_n < 1/n$ et $B_n \subset W \cap A_1^c \cap \dots \cap A_n^c$; alors $\bigcap_n B_n = \{x\}$ où $x \in W \cap A_1^c \cap A_2^c \cap \cdots$ contredisant l'hypothèse que $W = A_1 \cup A_2 \cup \cdots$)

REMARQUE. Cet exercice est un cas particulier d'un énoncé plus général appelé le théorème de Baire valable dans un espace métrique complet quelconque (Newman [31, p. 52], Dieudonné [13, tome 2, p. 81]).

4.3 THÉORÈME DES FONCTIONS IMPLICITES

Nous démontrons un théorème sur l'existence des fonctions implicites, signalé préalablement dans le paragraphe 4.1.2, comme conséquence du théorème d'inversion locale 4.2.2. Les notations sont celles fixées dans le paragraphe 4.1.3.

4.3.1 Proposition (*théorème des fonctions implicites*)

Soit $F : \Omega \to \mathbb{R}^M$ une fonction de classe C^k, $1 \leqslant k \leqslant \infty$, définie dans une partie ouverte Ω de $\mathbb{R}^{N+M}$; supposons que $(a;b) \in \Omega$ est tel que

$$F(a;b) = c\,, \quad \det F'_u(a;b) \neq 0\,.$$

Alors, il existe un voisinage ouvert U de $(a;b)$, $U \subset \Omega$, une boule ouverte $V \subset \mathbb{R}^N$, centrée à a, et une unique fonction continue $f : V \to \mathbb{R}^M$ tels que

$$f(a) = b\,; \quad F\big(x; f(x)\big) = c\,,\ x \in V \tag{4.15}$$

$$\big\{(x;u) \in U : F(x;u) = c\big\} = \big\{(x;u) : x \in V\,,\ u = f(x)\big\}\,. \tag{4.16}$$

La fonction f est de classe C^k et $f'(x)$ est donné par la formule (4.11).

DÉMONSTRATION

Sans perte de généralité, nous pouvons supposer que $c = 0 \in \mathbb{R}^M$ (en remplaçant F par $F - c$ si nécessaire). Posons

$$\varphi(x;u) = \big(x\,;\, F(x;u)\big)\,, \quad (x;u) \in \Omega\,;$$

alors

$$\varphi(a;b) = \big(a\,;\, F(a;b)\big) = (a;0)$$

$$\varphi'(a;b) = \begin{bmatrix} I_N & 0 \\ F'_x(a;b) & F'_u(a;b) \end{bmatrix}$$

où I_N est la matrice identité de type (N,N) et 0 est la matrice nulle de type (N,M). On vérifie que

$$\det \varphi'(a;b) = \det F'_u(a;b) \neq 0\,.$$

En appliquant le théorème d'inversion locale sur φ au point $(a;b)$, nous obtenons un voisinage ouvert $U \subset \Omega$ de $(a;b)$ et un voisinage ouvert U' de $(a;0)$, $U' \subset \mathbb{R}^{N+M}$, tels que φ est un C^k-difféomorphisme de U sur U'. De plus, nous choisissons U' de la forme $V \times W$ où V est une boule ouverte, centrée en a, dans $\mathbb{R}^N$, et W un voisinage ouvert de 0 dans $\mathbb{R}^M$; ceci est possible, quitte à prendre, après coup, U comme $\varphi^{-1}(U')$.

Posons $\psi : U' \to U$ pour la fonction (de classe C^k qui est inverse de $\varphi : U \to U'$; alors, compte tenu de la forme de φ, on aura $\psi(a;0) = (a;b)$ et

$$\psi(x;u) = \big(x; G(x;u)\big)\,, \quad (x;u) \in U' = V \times W$$

où $G : U' \to \mathbb{R}^M$ et telle que $G(a;0) = b$ et G est de classe C^k ; de plus

$$F\big(x; G(x;u)\big) = u\,, \quad (x;u) \in V \times W$$

car, si $(x;u) \in V \times W$,

$$(x;u) = \varphi\big(\psi(x;u)\big) = \varphi\big(x; G(x;u)\big) = \big(x\,;\, F\big(x; G(x;u)\big)\big)\,.$$

Définissons $f(x) = G(x;0)$, $x \in V$; alors

$$f(a) = G(a;0) = b\,, \quad F\big(x; f(x)\big) = F\big(x; G(x;0)\big) = 0\,,\ x \in V\,.$$

Ceci établit (4.15). Si $(x;u) \in U$ et $F(x;u) = 0$ alors

$$\varphi(x;u) = (x;0) \in V \times W$$

de sorte que $x \in V$; de plus $u = f(x)$ puisque

$$(x;u) = \psi(x;0) = \big(x; G(x;0)\big) = \big(x; f(x)\big)\,.$$

Inversement, si $x \in V$ et $u = f(x)$ on a

$$F(x;u) = F\big(x; f(x)\big) = F\big(x; G(x;0)\big) = 0$$
$$\big(x; f(x)\big) = \big(x; G(x;0)\big) = \psi(x;0) \in U$$

ce qui établit la formule (4.16) complètement. Puisque G est de classe C^k, $f : V \to \mathbb{R}^M$ est de classe C^k car $f(x) = G(x;0)$.

Démontrons maintenant l'unicité de f ; c'est ici que nous utilisons le fait que le voisinage ouvert V de $a \in \mathbb{R}^N$ est connexe (car V est une boule). Soit $g : V \to \mathbb{R}^M$ une fonction continue telle que

$$g(a) = b\,, \quad F\big(x; g(x)\big) = 0\,,\ x \in V\,.$$

Posons

$$E = \big\{x \in V : f(x) = g(x)\big\}\ ;$$

il est clair que E est non vide car $a \in E$, et que E est fermé car $E = (f-g)^{-1}(0)$; démontrons que E est aussi ouvert. En effet, soit $\xi \in E$; alors $\xi \in V$ et $F\big(\xi; g(\xi)\big) = F\big(\xi; f(\xi)\big) = 0$ de sorte que $\big(\xi; g(\xi)\big) \in U$ selon (4.16) ; puisque $x \mapsto \big(x; g(x)\big)$ est continue lorsque $x \in V$, il existe un voisinage ouvert B de ξ, $B \subset V$, tel que $\big(x; g(x)\big) \in U$ pour $x \in B$; de nouveau, selon (4.16), puisque $F\big(x; g(x)\big) = 0$, $x \in B$, on aura que $g(x) = f(x)$, $x \in B$, de sorte que $B \subset E$. Puisque V est connexe on conclut que $V = E$, c'est-à-dire $g = f$. ♦

4.3.2 Remarques

Il est important de saisir le sens pratique de la proposition 4.3.1 sur l'existence des fonctions implicites. Dans l'énoncé donné, on a considéré une fonction F de $N+M$ variables réelles, $x_1, \ldots, x_N, u_1, \ldots, u_M$, à valeurs dans $\mathbb{R}^M$; en pratique, les $N+M$ variables ne se présentent pas en deux blocs séparés comme indiqué ; cette présentation est choisie (par la plupart des auteurs) pour avoir un énoncé clair et une demonstration transparente. Pratiquement, on doit traiter une fonction F de n variables réelles $y_1, \ldots, y_n$ (variant dans une partie ouverte de $\mathbb{R}^n$) à valeurs dans $\mathbb{R}^M$ avec $M < n$; en supposant que F est de classe C^1 on peut étudier $F'(\alpha)$ pour un point α fixé ; si $F = (F_1, \ldots, F_M)$, ceci revient à étudier la matrice $[\mathrm{D}_j F_i]$ de type (M, n). Evidemment, le rang de cette matrice est $r \leqslant M$; supposons que $r = M$. Alors on sait que tous les mineurs d'ordre $> M$ sont nuls et qu'il existe au moins un mineur d'ordre M différent de 0 ; ainsi, il existe $j_1 < j_2 < \ldots < j_M$ tels que au point α

$$\det\left[\mathrm{D}_{j_k} F_i\right] \neq 0\,.$$

En écrivant $u_1 = y_{j_1}, \ldots, u_M = y_{j_M}$ et $x_1, \ldots, x_N$, $N = n - M$, pour les autres variables restantes, on peut reconsidérer le problème dans le formalisme de la proposition 4.3.1. Ce qui se passe si le rang $r < M$ sera discuté dans la suite de ce chapitre.

Revenons à l'énoncé de la proposition 4.3.1 (appelée : théorème des fonctions implicites). Géométriquement, il dit que si

$$F(a;b) = c \quad \text{et} \quad \det F'_u(a;b) \neq 0$$

alors l'ensemble

$$E = \left\{(x;u) \in \Omega : F(x;u) = c\right\}$$

localement autour du point $(a;b)$ (vu seulement dans le voisinage ouvert U de (a ;b) mentionné dans l'énoncé) est exactement égal au graphe $\big(x, f(x)\big)$, $x \in V$ d'une fonction $f : V \to \mathbb{R}^M$. Pour $M = N = 1$, ceci veut dire que localement E est une « courbe » de la forme $u = f(x)$; pour $N = 2$, $M = 1$, E sera localement une surface de la forme $u = f(x_1, x_2)$. Ainsi, pour $M = N = 1$, l'ensemble E ne peut pas avoir un point « multiple » comme dans (i) de la figure 4.2 ; il aura plutôt une forme comme dans (ii) de la même figure.

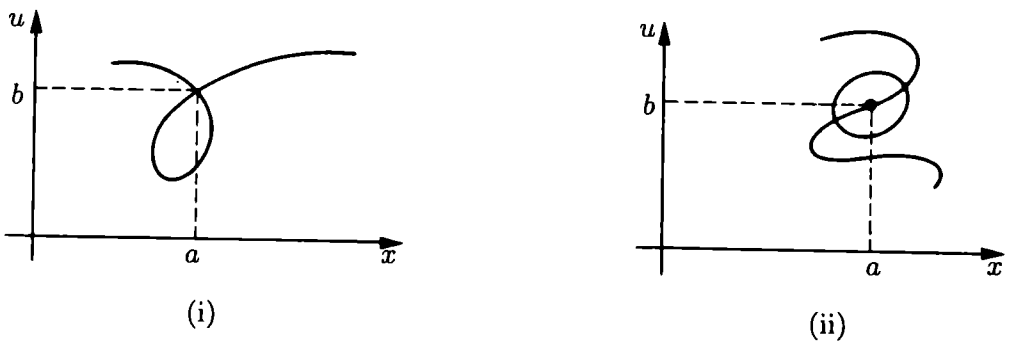


Fig. 4.2

Nous avons obtenu le théorème des fonctions implicites comme une conséquence du théorème d'inversion locale ; en fait, les deux théorèmes sont équivalents dans le sens qu'un des deux est une conséquence assez facilement déductible de l'autre, et nous avons déjà indiqué (§ 4.1.1) pourquoi le théorème d'inversion locale est un cas particulier du théorème des fonctions implicites. En utilisant d'autres méthodes de démonstration, on peut obtenir des énoncés plus généraux ; ainsi, notre appellation (qui est celle actuellement en usage) « théorème des fonctions implicites » ou « théorème d'inversion locale » est quelque peu constestable. Notons, en passant, que les premiers énoncés explicites de ces théorèmes sont dus à Cauchy autour des années 1820 (sous une hypothèse forte « d'analyticité » sur F) et à Dini (en 1877-78 dans son *Analisi infinitesimale*, vol. 1) dans une forme équivalente à celle donnée ici ; à partir de 1900, ces théorèmes, plus ou moins dans leur forme présentée ici, sont donnés et démontrés dans un grand nombres de cours d'analyse écrits par des auteurs connus comme Jordan, Goursat et bien d'autres. On a l'impression qu'avant Cauchy le problème d'*existence* de ces fonctions implicites n'était pas mis en question ; une équation $F(x;u) = 0$ définissait *évidemment* une fonction $u = f(x)$ et les discussions se bornaient aux calculs des dérivées de f, un peu comme dans le paragraphe 4.1.3.

4.3.3 Exercices

[1] Considérer l'équation $F(x,y) = 0$ autour du point $(0,0)$ dans chacun des trois exemples suivants :

(i) $F(x,y) = y^2 - x^2$;
(ii) $F(x,y) = y^3 - x$;
(iii) $F(x,y) = y^3 - x^4$.

Vérifier que dans chaque cas $F(0,0) = 0$, $F'_y(0,0) = 0$; l'ensemble $E = \{(x,y) \in \mathbb{R}^2 : F(x,y) = 0\}$ n'est pas le graphe d'une fonction $y = f(x)$ (ou $x = f(y)$) localement autour de $(0,0)$ dans le cas (i) ; E est de la forme $y = f(x)$ dans le cas (ii) mais f n'est pas dérivable au point $x = 0$; dans (iii), E est de la forme $y = f(x)$ avec f de classe C^1 partout.
(Cet exercice montre, entre autres, que la condition $\det F'_u(a;b) \neq 0$ dans la proposition 4.3.1 n'est pas une condition nécessaire pour la validité de sa conclusion.)

[2] Vérifier, grâce à la proposition 4.3.1, que dans chacun des exercices 2, 3, 4, 5 et 7 du paragraphe 4.1.4 les fonctions dont l'existence est prise comme hypothèse existent effectivement.

[3] Démontrer que l'équation $x^{2n+1} + y^{2n+1} + z^{2n+1} - \alpha xyz = 0$ (avec $\alpha \neq 0$, n entier $\geqslant 1$) possède une unique solution de classe C^∞ de la forme $z = f(x,y)$ définie dans un voisinage ouvert de $(1,-1)$ avec $f(1,-1) = 0$. Calculer $\nabla f(1,-1)$.

4 Peut-on représenter la « courbe » d'équation $x^2 + y + \sin(xy) = 1$ sous la forme $y = f(x)$ autour de $x = 1$, $y = 0$? Ecrire l'équation de la tangente à cette courbe à $(1, 0)$.

5 Donner des conditions suffisantes pour que le système d'équations suivant possède une solution de la forme $y = \varphi(x)$, $z = \psi(x)$ autour de $(x, y, z) = (1, 1, 1)$:

$$f(xy) + g(yz) = 0\,, \quad g(xy) + f(yz) = 0$$

où f, g sont des fonctions réelles de classe C^1 d'une variable réelle. Calculer l'angle d'intersection des tangentes aux graphes de φ et ψ au point $(1, 1)$.

6 Démontrer que l'on peut résoudre les équations $u^2 + v^2 = \ln x$, $u^3 + 2v^3 = \ln y$ dans un voisinage ouvert de $(u, v) = (1, 1)$ dans la forme $u = f(x, y)$, $v = g(x, y)$; calculer $F'(x, y)$ si $F = (f, g)$.

7 Supposons que (x_0, y_0, z_0, u_0) soit une solution du système de trois équations :

$$f(x) + f(y) + f(z) = F(u)$$
$$g(x) + g(y) + g(z) = G(u)$$
$$h(x) + h(y) + h(z) = H(u)$$

avec f, g, h, F, G, H de classe C^1 sur $\mathbb{R}$.

(a) Donner une condition suffisante pour pouvoir résoudre x, y, z en fonction de u dans un voisinage de (x_0, y_0, z_0, u_0).
(b) Appliquer cette condition au cas $f(x) = x$, $g(x) = x^2$, $h(x) = x^3$.
(c) Généraliser (a) et (b) au cas de n équations de forme analogue.

4.4 THÉORÈME DU RANG CONSTANT

4.4.1 Introduction

Dans cette section, nous discutons la nature des ensembles $f^{-1}(b)$ et $f(\Omega)$ pour une fonction $f : \Omega \to \mathbb{R}^m$, Ω étant une partie ouverte non vide de $\mathbb{R}^n$; nous supposerons que f est au moins de classe C^1 et que $\operatorname{rang} f'(x) = r$ est *constant* pour tout $x \in \Omega$. Nous dirons alors que f est de ***rang constant*** dans Ω.

Il s'agit donc d'une étude des solutions du système d'équations

$$f_i(x_1, \ldots, x_n) = y_i\,, \quad 1 \leqslant i \leqslant m\,, \ (x_1, \ldots, x_n) \in \Omega\,,$$

où nous avons écrit $f = (f_1, \ldots, f_m)$ pour la fonction de rang constant f.

Notons que *si* $f \in C^1(\Omega; \mathbb{R}^m)$, $a \in \Omega$, *et* $\operatorname{rang} f'(a) = r$, *alors il existe un voisinage ouvert* U *de* a, $U \subset \Omega$, *tel que* $\operatorname{rang} f'(x) \geqslant r$ *pour* $x \in U$.

En effet, $\operatorname{rang} f'(a) = r$ entraîne que la matrice $[\mathrm{D}_j f_i(a)]$ possède un mineur d'ordre r

$$\det[\mathrm{D}_{j_\beta} f_{i_\alpha}(x)]\,, \quad 1 \leqslant i_1, \ldots, i_r \leqslant m\,,\ 1 \leqslant j_1, \ldots, j_r \leqslant n$$

non nul pour $x = a$, ce qui entraîne que le même mineur est non nul pour x parcourant un voisinage ouvert U de a, $U \subset \Omega$, à cause de la continuité des fonctions $\mathrm{D}_j f_i$. Ainsi, $\operatorname{rang} f'(x) \geqslant r$ pour $x \in U$.

Pour simplifier l'écriture, nous supposerons que le mineur non nul d'ordre r en question est donné par $i_1 = 1, \ldots, i_r = r$ et $j_1 = 1, \ldots, j_r = r$. Ceci revient à renommer les variables $x_{j_1}, \ldots, x_{j_r}$ comme $x_1, \ldots, x_r$ et les fonctions $f_{i_1}, \ldots, f_{i_r}$ comme $f_1, \ldots, f_r$, et ne diminue en rien la généralité de la discussion. Donc nous supposerons que

$$\det[\mathrm{D}_j f_i(x)] \neq 0\,, \quad 1 \leqslant i, j \leqslant r\,,\ x \in \Omega\,. \tag{4.17}$$

La condition (4.17) garantit que $\operatorname{rang} f'(x) \geqslant r$ pour $x \in \Omega$; pour avoir $\operatorname{rang} f'(x) = r$, $x \in \Omega$, nous supposerons en plus de (4.17) que tous les mineurs d'ordre $> r$ de la matrice canonique de $f'(x)$ sont nuls pour $x \in \Omega$.

L'exemple le plus simple d'une application θ de rang constant r est le suivant : soit $\theta : \mathbb{R}^n \to \mathbb{R}^m$ définie par

$$\theta(x_1, \ldots, x_n) = (x_1, \ldots, x_r, 0, \ldots, 0)$$

avec une interprétation évidente si $r = 0$ ou m ; ici la matrice canonique de $\theta'(x)$ est de la forme

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

où I_r est la matrice identité de type (r, r), d'où $\operatorname{rang} \theta'(x) = r$ pour tout x. Le théorème du rang constant dit que toute application du rang constant r s'obtient de θ par des changements de variables appropriés.

Le cas de $r = 0$ est trivial ; on aura alors $\mathrm{D}_j f_i(x) = 0$ partout dans Ω, pour tout i, j ce qui équivaut à avoir les f_i constantes dans Ω (supposé connexe). Donc nous supposerons que $1 \leqslant r \leqslant \min(m, n)$.

Il est commode d'introduire ici la terminologie suivante : un pavé ouvert contenant un point $a \in \mathbb{R}^k$ s'appellera un ***voisinage rectangulaire*** de a. Donc si $a = (a_1, \ldots, a_k)$, un voisinage rectangulaire de a est un pavé $P = I_1 \times \cdots \times I_k$, chaque I_j étant un intervalle ouvert de $\mathbb{R}$, tel que $a \in P$.

4.4.2 Proposition (*théorème du rang constant*)

Soit $f : \Omega \to \mathbb{R}^m$ une application de classe C^1, définie dans une partie ouverte non vide Ω de $\mathbb{R}^n$, de rang constant $r \geqslant 1$. Soit $a \in \Omega$ et $f(a) = b$. Alors il existe :

- *un voisinage ouvert U de a, $U \subset \Omega$ et un C^1-difféomorphisme $u : U \to P$, P étant un voisinage rectangulaire de $0 \in \mathbb{R}^n$,*

- *un voisinage ouvert V de $b \in \mathbb{R}^m$ et un C^1-difféomorphisme $v : V \to Q$, Q étant un voisinage rectangulaire de $0 \in \mathbb{R}^m$,*

tels que

$$v \circ f \circ u^{-1}(x_1, \ldots, x_n) = (x_1, \ldots, x_r, 0, \ldots, 0)\,.$$

Si f est de classe C^p alors u, v seront des C^p-difféomorphismes.

DÉMONSTRATION

Supposons d'abord que $a = 0 \in \mathbb{R}^n$, $b = 0 \in \mathbb{R}^m$ et que f satisfait (4.17). Posons

$$\varphi(x) = \big(f_1(x), \ldots, f_r(x), x_{r+1}, \ldots, x_n\big)\,.$$

Alors $\varphi : \Omega \to \mathbb{R}^n$ est tel que

$$\varphi'(x) = \left[\begin{array}{c|c} \mathrm{D}_j f_i(x) & * \, * \, * \\ 1 \leqslant i, j \leqslant r & \\ \hline 0 & I_{n-r} \end{array}\right]$$

ce qui donne $\operatorname{rang} \varphi'(x) = n$, $x \in \Omega$. Selon le théorème d'inversion locale il existe un voisinage ouvert U_0 de 0, $U_0 \subset \Omega$ et un domaine $\widetilde{U}_0 \subset \mathbb{R}^n$ tels que φ est un C^1-difféomorphisme de U_0 sur $\widetilde{U}_0$. Comme $\varphi(0) = 0 \in \widetilde{U}_0$, on peut choisir un voisinage rectangulaire P de 0 tel que $P \subset \widetilde{U}_0$; prenons $U = \varphi^{-1}(P)$ et posons $u = \varphi|_U$. Alors $u : U \to P$ est un C^1-difféomorphisme. On vérifie que

$$f\big(u^{-1}(x)\big) = \big(x_1, \ldots, x_r, g_{r+1}(x), \ldots, g_m(x)\big) \tag{4.18}$$

où $x = (x_1, \ldots, x_n)$ et $g_{r+1}, \ldots, g_m$ sont des fonctions de classe C^1 dans P ; c'est une conséquence immédiate de la forme de $u(x) = \varphi(x)$, $x \in U$.

Maintenant, la formule

$$\big(f \circ u^{-1}\big)'(x) = f'\big(u^{-1}(x)\big)\big(u^{-1}\big)'(x)\,, \quad x \in P\,,$$

entraîne que $\operatorname{rang}\big(f \circ u^{-1}\big)'(x) = r$, $x \in P$, car $\big(u^{-1}\big)'(x)$ est inversible et $\operatorname{rang} f'(y) = r$ pour tout $y \in \Omega$. Or, de (4.18), on aura

$$\big(f \circ u^{-1}\big)'(x) = \left[\begin{array}{c|c} I_r & 0 \\ \hline * \, * \, * & \mathrm{D}_j g_i(x) \\ & r+1 \leqslant i, j \leqslant m \end{array}\right]\,.$$

Le rang de cette dernière matrice étant r on doit avoir que $\mathrm{D}_j g_i(x) = 0$, $x \in P$, pour tous les i, j entre $r+1$ et m. Ceci entraîne que les g_i ne dépendent pas des coordonnées $x_{r+1}, \ldots, x_m$; autrement dit,

$$g_i(x) = h_i(x_1, \ldots, x_r)\,, \quad x \in P\,,\ r+1 \leqslant i \leqslant m\,.$$

Posons maintenant

$$\psi(y_1, \ldots, y_m) = \big(y_1, \ldots, y_r, y_{r+1} - h_{r+1}(y_1, \ldots, y_r), \ldots, y_m - h_m(y_1, \ldots, y_r)\big) .$$

En écrivant $P = P^r \times P^{n-r}$, P^r étant un voisinage rectangulaire de $0 \in \mathbb{R}^r$ et P^{n-r} un voisinage rectangulaire de $0 \in \mathbb{R}^{n-r}$, on voit que

$$\psi : P^r \times \mathbb{R}^{m-r} \longrightarrow P^r \times \mathbb{R}^{m-r} .$$

On vérifie que $\psi(0) = 0$, ψ est une bijection de classe C^1 et que $\psi'(y)$ est inversible pour tout y dans son domaine de définition. En fait,

$$\psi^{-1}(y_1, \ldots, y_m) = \big(y_1, \ldots, y_r, y_{r+1} + h_{r+1}(y_1, \ldots, y_r), \ldots, y_m + h_m(y_1, \ldots, y_r)\big)$$

$$\psi'(y) = \left[\begin{array}{c|c} I_r & 0 \\ \hline * * * & I_{m-r} \end{array}\right]$$

ce qui permet de faire les vérifications indiquées.

Notons maintenant que, d'après (4.18), $f \circ u^{-1}$ est une application de P dans $V = P^r \times \mathbb{R}^{m-r}$, un voisinage ouvert de $0 \in \mathbb{R}^m$; V est aussi un voisinage rectangulaire de $0 \in \mathbb{R}^m$. Un calcul simple montre que, si $x = (x_1, \ldots, x_n) \in P$, on a

$$\psi\Big(f\big(u^{-1}(x)\big)\Big) = (x_1, \ldots, x_r, 0, \ldots, 0)$$

en utilisant (4.18), le fait que les $g_i = h_i$ et la définition de ψ. En posant $Q = V$ et v pour ψ on aura $v : V \to Q$ tel que u, U, P, v, V, Q satisfont tout ce qui est affirmé dans l'énoncé de la proposition 4.4.2. En plus, si f est de classe C^p, on vérifie immédiatement que u, v seront de classe C^p.

Passons maintenant au cas général de $f(a) = b$. Il est clair qu'il existe des applications affines $\sigma : \mathbb{R}^m \to \mathbb{R}^m$ et $\tau : \mathbb{R}^n \to \mathbb{R}^n$ telles que $\sigma(b) = 0$, $\tau(0) = a$ et telles que $\widetilde{f} = \sigma \circ f \circ \tau$ est une application de $\widetilde{\Omega} = \tau^{-1}(\Omega)$ dans $\mathbb{R}^m$ de rang constant r avec $\widetilde{f}(0) = 0$ et $\widetilde{f}$ satisfaisant la condition imposée par (4.17). Il suffit de prendre $\tau(x) = Ax + a$, $\sigma(y) = B(y - b)$ où A, B sont des matrices qui permutent les coordonnées $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_m)$ de manières appropriées ; alors

$$\widetilde{f}'(x) = \sigma'\big(f \circ \tau(x)\big) f'\big(\tau(x)\big) \tau'(x) = B f'\big(\tau(x)\big) A$$

montre que A, B peuvent être choisies de telle manière que $\widetilde{f}'(x)$ satisfait la condition imposée par (4.17). Si $\widetilde{u}$, $\widetilde{U}$, $\widetilde{P}$, $\widetilde{v}$, $\widetilde{V}$, $\widetilde{Q}$ sont choisis pour $\widetilde{f}$ comme indiqué dans la proposition, alors la formule

$$\widetilde{v} \circ \widetilde{f} \circ \widetilde{u}^{-1} = \widetilde{v} \circ \sigma \circ f \circ \tau \circ \widetilde{u}^{-1}$$

montre que $v = \widetilde{v} \circ \sigma$, $u = \widetilde{u} \circ \tau^{-1}$, $U = \tau^{-1}\widetilde{U}$, $P = \widetilde{P}$, $V = \sigma\widetilde{V}$, $Q = \widetilde{Q}$ ont les propriétés annoncées. ♦

REMARQUE. Notons que $f(U) \subset V$ car $v \circ f \circ u^{-1}(P) \subset Q$ d'où $v \circ f(U) \subset Q$ et $f(U) \subset v^{-1}(Q) = V$.

4.4.3 Remarques et définitions

Pour pouvoir formuler succinctement certaines conclusions importantes du théorème du rang constant, il est commode d'introduire quelques définitions.

Une application (de classe C^p), $f : \Omega \to \mathbb{R}^m$, Ω étant une partie ouverte de $\mathbb{R}^n$, du rang constant n, s'appelle une ***immersion*** (de classe C^p) de Ω dans $\mathbb{R}^m$; il est clair que m est alors $\geqslant n$; si l'immersion (de classe C^p) f est un homéomorphisme entre Ω et $f(\Omega) \subset \mathbb{R}^m$, alors f s'appelle un ***plongement*** (de classe C^p) de Ω dans $\mathbb{R}^m$. Pour que f soit un plongement (de classe C^p) de $\Omega \subset \mathbb{R}^n$ dans $\mathbb{R}^m$ il faut et il suffit donc que f soit une immersion (de classe C^p) de Ω dans $\mathbb{R}^m$ qui est injective et qui est telle que $f^{-1} : f(\Omega) \to \Omega$ est continue. L'exercice 1 (§ 4.2.10) montre que la dernière précision n'est pas inutile.

Soit $E \subset \mathbb{R}^n$, $E \neq \emptyset$; on dit que E est une ***variété différentielle*** de dimension k ($1 \leqslant k \leqslant n$) de classe C^p contenue dans $\mathbb{R}^n$ (ou une ***sous-variété différentielle*** de $\mathbb{R}^n$ de dimension k, de classe C^p), si pour tout point $a \in E$ on peut déterminer un voisinage ouvert U de a tel que $U \cap E = \psi(V)$ où V est une partie ouverte de $\mathbb{R}^k$, et $\psi : V \to \mathbb{R}^n$ est un plongement de classe C^p.

Les sous-variétés différentielles de $\mathbb{R}^n$ sont des parties qui reviennent très souvent dans les considérations géométriques et analytiques. Les sous-variétés différentielles connexes de dimension 1 dans $\mathbb{R}^n$ (de classe C^p) sont ce que nous appellerons les ***courbes régulières simples*** (de classe C^p) dans $\mathbb{R}^n$; l'adjectif « régulier » indique que la courbe possède une tangente en chacun de ses points, variant continûment le long de la courbe ; l'adjectif « simple » indique que la courbe ne possède pas des points multiples où la courbe se croise elle-même ; ces considérations seront reprises dans la suite. Les sous-variétés différentielles de dimension 2 dans $\mathbb{R}^n$, $n \geqslant 2$ (de classe C^p) seront appelées les ***surfaces régulières*** dans $\mathbb{R}^n$; l'adjectif « régulier » veut dire ici, intuitivement, que la surface possède un plan tangent à chacun de ses points, variant continûment d'un point de la surface à un autre ; les surfaces, par définition, ne possèdent pas de « points multiples » même si le mot « surface » dans son usage quotidien permet des objets plus généraux. Les surfaces régulières dans $\mathbb{R}^3$ seront discutées plus longuement dans la suite.

Il est peut-être utile de remarquer, de nouveau, que même si les définitions de courbes régulières simples et surfaces régulières reflètent bien nos idées intuitives des courbes et des surfaces, ces définitions ne contiennent pas tous les exemples intéressants rencontrés de ces derniers objets. Ainsi, une « courbe » aussi simple que celle de la forme d'un 6 dans $\mathbb{R}^2$ n'est pas une variété différentielle de dimension 1, à cause de l'existence d'un point où la courbe revient sur elle-même.

L'idée intuitive derrière la définition d'une sous-variété de dimension k, $E \subset \mathbb{R}^n$, est que tout point $a \in E$ possède un voisinage ouvert U tel que $U \cap E$ peut être ***paramétré*** par un ***paramètre*** de dimension k parcourant une partie ouverte $V \subset \mathbb{R}^k$, le paramétrage étant effectué par une fonction $\psi : V \to \mathbb{R}^n$ convenable dans le sens que ψ est un plongement (de classe C^p). On dit aussi qu'une sous-variété différentielle de dimension k, dans $\mathbb{R}^n$, possède une ***paramétrisation locale*** de dimension k. Il est utile d'appeler (ψ, V) une paramétrisation locale (de $U \cap E$) autour de $a \in E$.

Une application (de classe C^p) $f : \Omega \to \mathbb{R}^m$, Ω étant une partie ouverte de $\mathbb{R}^n$, de rang constant m s'appelle une ***submersion*** (de classe C^p) de Ω dans $\mathbb{R}^m$; il est clair que m est alors $\leqslant n$.

Un exemple typique d'une submersion est l'application $\theta : \mathbb{R}^n \to \mathbb{R}^m$, $m < n$, avec $\theta(x_1, \ldots, x_n) = (x_1, \ldots, x_m)$. L'exemple typique d'une immersion est l'application $\theta : \mathbb{R}^n \to \mathbb{R}^m$, $m > n$, avec $\theta(x_1, \ldots, x_n) = (x_1, \ldots, x_n, 0, \ldots, 0)$, mais celle-ci est aussi un plongement.

Une immersion est, localement, comme une inclusion, et une submersion est, localement, comme une projection. Ces remarques intuitives sont rendues précises par le théorème du rang constant (avec $r = m$ pour une submersion et $r = n$ pour une immersion).

Ce que nous avons appelé une sous-variété différentielle de $\mathbb{R}^n$ s'appelle souvent une ***sous-variété différentielle sans bord*** dans $\mathbb{R}^n$. La notion d'une sous-variété différentielle avec bord (ou à bord) sera introduite dans la suite.

Il est utile de définir une ***sous-variété différentielle de dimension zéro*** dans $\mathbb{R}^n$ comme une partie $E \subset \mathbb{R}^n$ qui est ***discrète*** dans le sens que tout point $a \in E$ possède un voisinage ouvert U tel que $U \cap E = \{a\}$.

Notons aussi que, selon notre définition, une sous-variété différentielle de dimension n dans $\mathbb{R}^n$ est une partie ouverte non vide de $\mathbb{R}^n$ et vice versa ; il est utile de vérifier cette affirmation simple en détail.

Nous donnons maintenant une description différente des sous-variétés de $\mathbb{R}^n$, qui est souvent utile.

4.4.4 Proposition

Soit E une partie non vide de $\mathbb{R}^n$. Pour que E soit une sous-variété différentielle de dimension k, $1 \leqslant k \leqslant n$, de classe C^p, il faut et il suffit que E satisfasse à la condition suivante (C) : pour tout $a \in E$ il existe un voisinage ouvert U de a dans $\mathbb{R}^n$, une partie ouverte $V \subset \mathbb{R}^n$, un C^p-difféormorphisme $f : U \to V$ tels que

$$f(U \cap E) = \{y = (y_1, \ldots, y_n) \in V : y_{k+1} = \ldots = y_n = 0\}. \tag{4.19}$$

REMARQUE. Le cas $k = n$ de (4.19) doit être interprété comme la formule $f(U \cap E) = V$. En écrivant $f = (f_1, \ldots, f_n)$, la condition (4.19) dit que $x \in U \cap E$ si et seulement si

$$f_{k+1}(x) = 0, \ldots, f_n(x) = 0.$$

En remplaçant f par $f + b$ et V par $V + b$ avec $b = (b_1, \ldots, b_k, 0, \ldots, 0)$ approprié, on peut même exiger que $f(a) = 0$.

Il est utile d'appeler (f, U) une ***carte redressante*** de E au point a.

Dans les applications, il est plus commode de reformuler (4.19) d'une manière plus flexible. On dira que (f, U) est une carte redressante de E au point a s'il existe V, $f : U \to V \subset \mathbb{R}^n$, comme dans l'énoncé de la proposition 4.4.3, avec (4.19) remplacé par

$$f(U \cap E) = V \cap F \tag{4.20}$$

où F est un sous-espace vectoriel (quelconque) de dimension k dans $\mathbb{R}^n$. Dans (4.19)

$$F = \{y = (y_1, \ldots, y_n) \in \mathbb{R}^n : y_{k+1} = \ldots = y_n = 0\}.$$

La preuve formelle de l'équivalence de ces deux formulations est laissée en exercice.

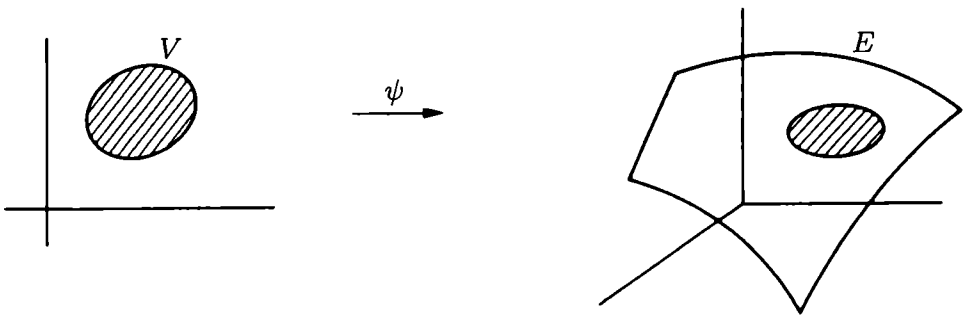


Fig. 4.3 (ψ, V) paramétrisation locale.

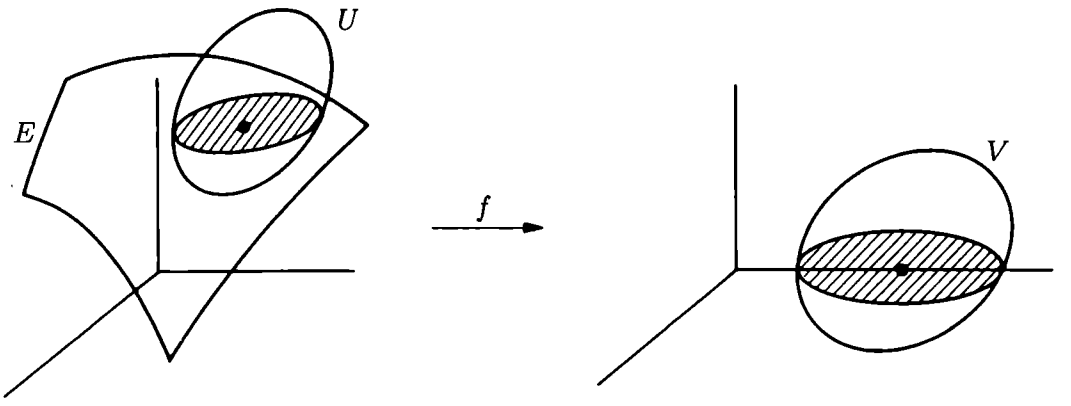


Fig. 4.4 (f, U) carte redressante.

DÉMONSTRATION

Supposons que E est une partie non vide de $\mathbb{R}^n$ satisfaisant à la condition (C) de la proposition avec f, U, V associés. Posons

$$W = \{y = (y_1, \ldots, y_k) \in \mathbb{R}^k : (y_1, \ldots, y_k, 0, \ldots, 0) \in V\}$$

$$\psi(y_1, \ldots, y_k) = f^{-1}(y_1, \ldots, y_k, 0, \ldots, 0), \quad (y_1, \ldots, y_k) \in W.$$

Puisque V est ouvert dans $\mathbb{R}^n$ et $f : U \to V$ est un C^p-difféomorphisme, on voit que W est ouvert dans $\mathbb{R}^k$ et $\psi : W \to \mathbb{R}^n$ est de classe C^p ; de plus, selon (4.19), $\psi(W) = U \cap E$ et $\psi : W \to U \cap E$ est un homéomorphisme. Il reste à voir que $\operatorname{rang} \psi'(y) = k$ pour $y \in W$ pour établir que (ψ, W) est une paramétrisation locale de E au point $a \in E$. En effet, $\operatorname{rang} \psi'(y) \leqslant k$, $y \in W$, car $\psi'(y)$ est une application linéaire de $\mathbb{R}^k$ dans $\mathbb{R}^n$, $k \leqslant n$; d'autre part, si $\pi : \mathbb{R}^n \to \mathbb{R}^k$ est donnée par $\pi(x_1, \ldots, x_n) = (x_1, \ldots, x_k)$, alors $\pi \circ f \circ \psi(y) = y$, $y \in W$, d'où

$$(\pi \circ f)' \psi'(y) = I_k\,, \quad y \in W\,,$$

ce qui donne $\operatorname{rang} \psi'(y) \geqslant k$, $y \in W$, établissant le résultat désiré.

Supposons maintenant que E est une sous-variété différentielle de dimension k (de classe C^p) dans $\mathbb{R}^n$. Si $a \in E$, soit (ψ, W) une paramétrisation locale autour de a ($\psi(W) = G \cap E$, G ouvert dans $\mathbb{R}^n$) choisie telle que $0 \in W$, $\psi(0) = a$ et telle que, si $\psi = (\psi_1, \ldots, \psi_n)$,

$$\det\left[\mathrm{D}_j \psi_i(y)\right] \neq 0\,, \quad 1 \leqslant i, j \leqslant k\,, \; y \in W\,,$$

ce qui est possible car $\operatorname{rang} \psi'(y) = k$, $y \in W \subset \mathbb{R}^k$. En écrivant $\mathbb{R}^n$ comme $\mathbb{R}^k \times \mathbb{R}^{n-k}$ et les éléments de $\mathbb{R}^n$ comme $(y; z)$, $y \in \mathbb{R}^k$, $z \in \mathbb{R}^{n-k}$, posons

$$g(y; z) = \psi(y) + (0; z)\,, \quad y \in W\,.$$

Alors $g : W \times \mathbb{R}^{n-k} \to \mathbb{R}^n$, $g(y; 0) = \psi(y)$ si $y \in W$ de sorte que

$$g\big(W \times \{0\}\big) = G \cap E\,.$$

Il est clair que g est de classe C^p et $g(0; 0) = a$; un calcul simple montre que

$$g'(y; z) = \left[\begin{array}{c|c} \mathrm{D}_j \psi_i(y) & 0 \\ 1 \leqslant i, j \leqslant k & \\ \hline * \, * \, * & I_{n-k} \end{array}\right]$$

de sorte que $\det g'(0; 0) \neq 0$. Selon le théorème d'inversion locale, il existe un voisinage ouvert V' de $0 = (0; 0)$ contenu dans $W \times \mathbb{R}^{n-k}$ et un voisinage ouvert U' de $a \in \mathbb{R}^n$ tels que g établit un C^p-difféomorphisme entre V' et U'. Si

$$W' = \big\{y \in W : (y; 0) \in V'\big\}$$

alors $W' \subset W$ est une partie ouverte non vide dans $\mathbb{R}^k$; ψ étant un homéomorphisme entre W et $\psi(W) = G \cap E$, on a que $\psi(W')$ est un sous-ensemble relativement ouvert de $\psi(W)$ de sorte que $\psi(W') = U \cap E$ où U est une partie ouverte de $\mathbb{R}^n$ contenue dans U'. Posons, maintenant, $V = g^{-1}(U)$ et $f = g^{-1}\big|_U$; il est alors clair que U est un voisinage ouvert de $a \in \mathbb{R}^n$, V un voisinage ouvert de $0 \in \mathbb{R}^n$, $f : U \to V$, un C^p-difféomorphisme tel que

$$f(U \cap E) = \big\{x \in V : x_{k+1} = \ldots = x_n = 0\big\}$$

où $x = (x_1, \ldots, x_n)$; la dernière relation résulte du fait que

$$g\big(V \cap \mathbb{R}^k \times \{0\}\big) = U \cap E$$

et $f = g^{-1}\big|_U$. ♦

4.4.5 Immersions

Soit $f : \Omega \to \mathbb{R}^m$ une immersion de classe C^p d'une partie ouverte Ω de $\mathbb{R}^n$ dans $\mathbb{R}^m$ (avec $1 \leqslant n < m$). Alors :

(i) pour chaque $b \in \mathbb{R}^m$, $f^{-1}(b)$ est soit vide, soit un sous-ensemble discret de Ω (c'est-à-dire une variété de dimension 0) ;

(ii) l'ensemble $f(\Omega) \subset \mathbb{R}^m$ a la structure suivante : si $a \in \Omega$, $b = f(a) \in f(\Omega)$, il existe un voisinage ouvert U de a, $U \subset \Omega$, et un voisinage ouvert V de b dans $\mathbb{R}^m$ tels que $f(U) \cap V$ est une sous-variété différentielle de dimension n dans $\mathbb{R}^m$ de classe C^p.

En effet, si $f^{-1}(b)$ est non vide, soit $a \in U$ tel que $f(a) = b$; soit u, U, v, V, P, Q comme dans le théorème du rang constant ; alors

$$v \circ f \circ u^{-1}(x_1, \ldots, x_n) = (x_1, \ldots, x_n, 0, \ldots, 0)\,.$$

On vérifie immédiatement que si $\alpha \in U$ et $f(\alpha) = b$, alors $\alpha = a$ de sorte que $U \cap f^{-1}(b) = \{a\}$; ceci établit (i). Pour (ii), posons $\psi = f \circ u^{-1}$; alors $\psi : P \to V \cap f(U)$ est une paramétrisation appropriée de $V \cap f(U)$ car, selon (4.18), si $x = (x_1, \ldots, x_n) \in P$,

$$\psi(x) = \big(x_1, \ldots, x_n,\ h_{n+1}(x), \ldots, h_m(x)\big)$$

(où les fonctions $h_{n+1}, \ldots, h_m$ sont de classe C^p et égales aux $g_{n+1}, \ldots, g_m$ de (4.18)) ; il est clair que $\operatorname{rang} \psi'(x) = n$, $x \in P$ et que ψ est un homéomorphisme de P sur $V \cap f(U)$. ♦

4.4.6 Submersions

Soit $f : \Omega \to \mathbb{R}^m$ une submersion de classe C^p d'une partie ouverte Ω de $\mathbb{R}^n$ dans $\mathbb{R}^m$ (avec $1 \leqslant m < n$). Alors :

(i) pour chaque $b \in \mathbb{R}^m$, $f^{-1}(b)$ est soit vide, soit une sous-variété différentielle de dimension $n - m$ dans $\mathbb{R}^n$ (contenue dans Ω) de classe C^p ;

(ii) f est une application ouverte ; en particulier, $f(\Omega)$ est ouverte dans $\mathbb{R}^m$.

DÉMONSTRATION

(i) Supposons, sans perte de généralité, que $b = 0$ (en passant à $f - b$ par exemple) et que $f^{-1}(0) \neq \emptyset$. Soit $a \in f^{-1}(0)$. Si u, U, P sont comme dans le théorème du rang constant, on a que

$$f\big(u^{-1}(y_1, \ldots, y_n)\big) = (y_1, \ldots, y_m) \tag{4.21}$$

selon (4.18). Si $E = f^{-1}(0)$, on voit que

$$u(U \cap E) = \{y = (y_1, \ldots, y_n) \in P : y_1 = y_2 = \ldots = y_m = 0\}$$

car $y \in u(U \cap E) \Leftrightarrow y \in P$ et $u^{-1}(y) \in f^{-1}(0) \Leftrightarrow y \in P$ et $f(u^{-1}(y)) = 0$. Donc $u(U \cap E) = P \cap F$ où

$$F = \{(y_1, \ldots, y_n) \in \mathbb{R}^n : y_1 = \ldots = y_m = 0\}$$

est un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $n-m$. Ceci établit (i) selon la proposition 4.4.4 et la remarque qui la suit.

(ii) Si $a \in \Omega$ et u, U, P sont comme auparavant, alors de (4.21), on conclut que

$$f(U) = f(u^{-1}(P)) = \pi(P)$$

où $\pi(y_1, \ldots, y_n) = (y_1, \ldots, y_m)$; puisque P est ouvert dans $\mathbb{R}^n$, $\pi(P)$ est ouvert dans $\mathbb{R}^m$. Ainsi, tout point $a \in \Omega$ possède un voisinage ouvert U tel que $f(U)$ est ouvert dans $\mathbb{R}^m$. On en conclut aisément que $f(V)$ est ouvert si V est ouvert et $V \subset \Omega$. ♦

4.4.7 Valeurs régulières

La propriété des submersions formulée dans (i) (§ 4.4.6) donne une méthode commode pour définir des sous-variétés différentielles dans $\mathbb{R}^N$. La terminologie suivante est généralement associée à cette méthode.

Soit $f : \Omega \to \mathbb{R}^M$ une application dérivable, Ω étant ouvert dans $\mathbb{R}^N$. Un point $b \in \mathbb{R}^M$ s'appelle une ***valeur régulière*** de f si, pour tout $a \in f^{-1}(b)$, $\operatorname{rang} f'(a) = M$ ou si $f^{-1}(b) = \emptyset$. Un point $b \in \mathbb{R}^M$ qui n'est pas une valeur régulière s'appelle une ***valeur singulière*** (ou une ***valeur critique***) de f.

Noter que cette terminologie concerne $f(\Omega) \subset \mathbb{R}^M$, l'ensemble des valeurs de f. En ce qui concerne les points de Ω on a une terminologie analogue. Un point $a \in \Omega$ s'appelle un ***point singulier*** pour f si $\operatorname{rang} f'(a) < M$; certains auteurs appellent un tel point a aussi un ***point critique*** de f, mais nous éviterons cette appellation car, pour nous, un point critique de f est un point a tel que $f'(a) = 0$ (§ 2.4.2) ; évidemment, un point critique a de f dans notre sens est un point singulier de f car alors $\operatorname{rang} f'(a) = 0$. Si $a \in \Omega$ n'est pas un point singulier de f, on dit que a est un ***point régulier*** de f.

Faisons quelques remarques évidentes à propos de cette terminologie. Si $M > N$, il est clair que tous les points de Ω sont des points singuliers pour f car, alors, $\operatorname{rang} f'(a) \leqslant N < M$ pour tout $a \in \Omega$. Notons aussi une étrangeté entraînée par cette terminologie (généralement acceptée) : une valeur singulière b pour f peut très bien être l'image de plusieurs points réguliers de f ; il suffit pour cela que $f^{-1}(b)$ contienne au moins un point singulier et d'autres points réguliers de f (Ex. 2, § 4.4.10). Par contre, si b est une valeur régulière pour f, alors $f^{-1}(b)$ est constitué exclusivement des points réguliers de f, à moins que $f^{-1}(b)$ ne soit pas vide.

Si la fonction $f : \Omega \to \mathbb{R}^M$, Ω ouvert dans $\mathbb{R}^N$, $N \geqslant M$, est de classe C^p, et b est une valeur régulière de f telle que $f^{-1}(b) \neq \emptyset$, alors $f^{-1}(b)$ est une sous-variété différentielle dans $\mathbb{R}^N$ de dimension $N - M$ et de classe C^p.

En effet, chaque point $a \in f^{-1}(b)$ possède un voisinage ouvert $U_a \subset \Omega$ dans lequel rang $f'(x) = M$, $x \in U_a$ (§ 4.4.1), d'où l'existence de l'ensemble ouvert $U = \bigcup(U_a) \subset \Omega$ tel que $f|_U \to \mathbb{R}^M$ est une submersion ; aussi, $f^{-1}(b) \subset U$. La proposition du paragraphe 4.4.6 donne alors le résultat énoncé. ♦

Mentionnons un théorème important (appelé le ***théorème de Sard***) concernant les points singuliers d'une fonction $f : \Omega \to \mathbb{R}^M$, Ω ouvert dans $\mathbb{R}^N$, de classe C^p.

Si $p \geqslant \max(N - M + 1, 1)$ et S l'ensemble des points singuliers de f, alors $f(S)$ est un ensemble de mesure zéro dans $\mathbb{R}^M$.

La notion d'ensemble de mesure zéro sera expliquée dans le chapitre 5 ; pour apprécier l'énoncé du théorème de Sard, il suffit de savoir pour l'immédiat qu'un ensemble de mesure zéro est d'intérieur vide. En considérant le cas particulier $N < M$ où Ω tout entier constitue l'ensemble des points singuliers de f, on conclut que *si $f \in C^1(\Omega; \mathbb{R}^M)$, $\Omega \subset \mathbb{R}^N$, $N < M$, Ω ouvert, alors $f(\Omega)$ est un ensemble de mesure zéro* (appelé aussi ***ensemble négligeable***) dans $\mathbb{R}^M$. En se spécialisant au cas d'une immersion f on voit que ceci entraîne que les sous-variétés différentielles de dimension $N < M$ dans $\mathbb{R}^M$ sont des ensembles négligeables. Ainsi, une courbe ou une surface (régulières) dans $\mathbb{R}^3$ forment des ensembles négligeables (c'est-à-dire de volume zéro), ce qui est intuitivement clair. Il est rassurant, en tout cas pour f de classe C^1, que l'on ne puisse pas avoir les courbes de Peano qui remplissent des volumes. La démonstration du théorème de Sard n'est pas facile ; on trouvera une démonstration dans Dieudonné [13, tome III, pp. 161 et 164].

Une conséquence « pratique » du théorème de Sard est qu'étant donné $f \in C^p(\Omega; \mathbb{R}^M)$, Ω ouvert dans $\mathbb{R}^N$, un choix « au hasard » d'un $b \in \mathbb{R}^M$ donne une valeur régulière pour f ; par conséquent, $f^{-1}(b)$ est une sous-variété différentielle de $\mathbb{R}^N$, si elle n'est pas vide, pour presque tout choix de $b \in \mathbb{R}^M$.

Notons en conclusion de ce paragraphe que $f^{-1}(b)$ peut être un sous-ensemble de $\mathbb{R}^N$ assez compliqué si b est une valeur singulière de f (Ex. 6 et 7, § 4.4.10).

4.4.8 Applications de rang r

Il s'agit ici du cas général d'une application $f : \Omega \to \mathbb{R}^m$, Ω ouvert dans $\mathbb{R}^n$, de classe C^p, de rang constant r ; le cas $r = n$ (immersion) était discuté dans le paragraphe 4.4.5 et le cas $r = m$ (submersion) dans le paragraphe 4.4.6. Les énoncés sont complètement analogues à ceux des paragraphes 4.4.5 et 4.4.6 ; leurs démonstrations aussi s'obtiennent du théorème du rang constant comme dans les cas $r = n$ ou $r = m$. Nous nous bornons, donc, ici seulement à la formulation des énoncés correspondants.

(i) Pour chaque $b \in \mathbb{R}^m$ *tel que* $f^{-1}(b) \neq \emptyset$, *l'ensemble* $f^{-1}(b)$ *est une sous-variété différentielle dans* $\mathbb{R}^n$ *de dimension* $n - r$ *et de classe* C^p.

(ii) L'ensemble $f(\Omega) \subset \mathbb{R}^m$ *a la structure suivante : si* $a \in \Omega$, $b = f(a) \in f(\Omega)$, *il existe un voisinage ouvert* U *de* a, $U \subset \Omega$, *et un voisinage ouvert* V *de* b *dans* $\mathbb{R}^m$ *tels que* $f(U) \cap V$ *est une sous-variété différentielle dans* $\mathbb{R}^m$ *de dimension* r *et de classe* C^p.

4.4.9 Dépendance fonctionnelle

Une question qui se pose souvent est la suivante : étant donné un système de fonctions $f_1, \ldots, f_m$ de n variables réelles, est-ce que ce système forme un système « indépendant » ou y a-t-il une « dépendance fonctionnelle » parmi les fonctions du système ? Donner une réponse précise et générale à cette question, après l'avoir bien entendu formulée de manière exacte, n'est pas une tâche facile. Nous nous bornons à faire deux observations utiles ; aussi, nous supposerons que les fonctions f_i sont de classe C^1 au moins, définies dans une partie ouverte non vide Ω de $\mathbb{R}^n$ de sorte que $f = (f_1, \ldots, f_m) : \Omega \to \mathbb{R}^m$ est de classe C^1 (au moins) ; supposons, de plus, que f est de rang constant r.

Si $r = m$, f est une submersion, donc une application ouverte ; ici, naturellement, $n \geqslant m$, c'est-à-dire, en termes simples, le nombre de variables indépendantes n, $(x_1, \ldots, x_n) \in \Omega$, est supérieur ou égal au nombre de fonctions m du système $f_1, \ldots, f_m$. Alors, $f(\Omega)$ est une partie ouverte non vide de $\mathbb{R}^m$; si $\varphi : \Omega' \to \mathbb{R}$ est une fonction quelconque définie dans une partie ouverte Ω' contenant $f(\Omega)$ et si $\varphi \circ f \equiv 0$, alors $\varphi \equiv 0$ sur $f(\Omega)$. On peut qualifier cette situation comme suit : $f_1, \ldots, f_m$ forme un système de fonctions ***fonctionnellement indépendant*** ; en effet, $f_1(x), \ldots, f_m(x)$ peuvent prendre des valeurs arbitraires remplissant la partie ouverte $f(\Omega)$ lorsque x parcourt Ω et aucune relation fonctionnelle non triviale existe entre les f_i.

Par contre, si $r < m$, il est naturel de dire que $f_1, \ldots, f_m$ est un système ***fonctionnellement dépendant*** pour la raison suivante : selon la démonstration du théorème du rang constant, pour tout $a \in \Omega$, il existe un voisinage ouvert U de a, un voisinage ouvert P de $0 \in \mathbb{R}^n$ et un homéomorphisme $u : U \to P$ tels que

$$f\big(u^{-1}(x)\big) = \big(x_1, \ldots, x_r\,, h_{r+1}(x_1, \ldots, x_r), \ldots, h_m(x_1, \ldots, x_r)\big)\,, \quad x \in P$$

ce qui est (4.18) avec l'identification des $g_{r+1}, \ldots, g_m$ avec les $h_{r+1}, \ldots, h_m$. Comme

$$f(x) = f\left(u^{-1}\big(u(x)\big)\right)\,, \quad x \in U\,,$$

on en conclut immédiatement que, pour $x \in U$,

$$\begin{aligned} f_{r+1}(x) &= h_{r+1}\big(f_1(x), \ldots, f_r(x)\big) \\ &\vdots \\ f_m(x) &= h_m\big(f_1(x), \ldots, f_r(x)\big)\,. \end{aligned}$$

Ainsi, si f est une application de classe C^1 de rang constant, on peut dire simplement :

$$r = m \Longrightarrow \text{indépendance fonctionnelle}$$
$$r < m \Longrightarrow \text{dépendance fonctionnelle}.$$

En fait, la conclusion subsiste même si f n'est pas de rang constant pourvu que l'on formule la dépendance fonctionnelle de manière un peu plus subtile et que l'on exige que f soit de classe C^p, $p \geqslant \max(n-m+1\,,\,1)$, pour pouvoir utiliser le théorème de Sard. Indiquons d'abord la définition exacte : $f = (f_1, \ldots, f_m)$ de classe C^p dans une partie ouverte non vide $\Omega \subset \mathbb{R}^n$ s'appelle un système ***fonctionnellement dépendant*** dans un sous-ensemble $S \subset \Omega$, s'il existe un ensemble ouvert $\Omega' \supset f(S)$ et une fonction $g \in C^1(\Omega')$ tels que $g \circ f \equiv 0$ sur S et $g^{-1}(0)$ est d'intérieur vide. On a alors le résultat suivant : supposons que $f = (f_1, \ldots, f_m) \in C^p(\Omega; \mathbb{R}^m)$, $p \geqslant \max(n-m+1\,,\,1)$, Ω ouvert non vide dans $\mathbb{R}^n$; alors f *est un système fonctionnellement dépendant dans chaque partie compacte contenue dans* Ω *si et seulement si* $\operatorname{rang} f'(x) < m$ *pour tout* $x \in \Omega$.

En effet, si pour un $a \in \Omega$, $\operatorname{rang} f'(a) = m$, alors on aura $\operatorname{rang} f'(x) = m$ pour x dans un voisinage ouvert de a d'où, selon une remarque précédente, f est un système fonctionnellement indépendant (c'est-à-dire contraire de fonctionnellement dépendant). Pour la réciproque, indiquons le raisonnement général sans entrer dans les détails ; si $\operatorname{rang} f'(x) < m$ pour tout $x \in \Omega$, $f(\Omega)$ est de mesure zéro dans $\mathbb{R}^m$ selon le théorème de Sard ; si $K \subset \Omega$ est compact, $f(K)$ sera aussi de mesure zéro dans $\mathbb{R}^m$, car $f(K) \subset f(\Omega)$, de sorte que $f(K)$ est d'intérieur vide ; mais $F = f(K)$ étant une partie fermée dans $\mathbb{R}^m$, il existe (ceci est à démontrer) une fonction $g \in C^\infty(\mathbb{R}^m)$ telle que $F = g^{-1}(0)$; alors $g\big(f(x)\big) = 0$ pour tout $x \in K$ et $g^{-1}(0) = f(K)$ est d'intérieur vide. On trouvera les détails de ce raisonnement dans le livre de Narasimhan [29, p. 28].

Une application particulière de la discussion de ce paragraphe est dans les considérations des coordonnées curvilignes dans $\mathbb{R}^n$ introduites dans la section 3.8. Un système $f = (f_1, \ldots, f_n)$ de fonctions de classe C^1 dans une partie ouverte Ω' de $\mathbb{R}^n$ définit des ***coordonnées curvilignes*** pour $\Omega = f(\Omega')$ si $f : \Omega' \to \Omega$ est un C^1-difféomorphisme (donc Ω est aussi ouvert dans $\mathbb{R}^n$) ; si $f(u) = x$, $(u_1, \ldots, u_n) = u$ sont les coordonnées curvilignes de $x = (x_1, \ldots, x_n)$ définies par f.

Dans la terminologie de ce paragraphe, le système f est fonctionnellement indépendant. La terminologie donnée correspond à celle des présentations classiques (comme coordonnées polaires dans $\mathbb{R}^2$, coordonnées sphériques dans $\mathbb{R}^3$, etc., discutées dans la section 3.8) ; dans les écrits modernes, on a tendance à dire que $f_1(u), \ldots, f_n(u)$ sont des ***coordonnées locales*** du point $u \in \Omega'$. En tout cas, il y a accord absolu sur le fait que $f : \Omega' \to \Omega$ est un C^1-difféomorphisme ; si $f(u) = x$, on peut dire que $x = (x_1, \ldots, x_n)$ sont des coordonnées locales de u, et $u = (u_1, \ldots, u_n)$ sont des coordonnées curvilignes de x. C'est un peu redondant mais ces terminologies existent.

4.4.10 Exercices

[1] Vérifier que les expressions suivantes définissent des fonctions f de classe C^∞ dans certains sous-ensembles ouverts Ω de $\mathbb{R}^3$ et que 0 est une valeur régulière de chacune d'elles de sorte que $f^{-1}(0)$ est une surface régulière (bien connue) dans $\mathbb{R}^3$ de classe C^∞ :

(i) $x^2 + y^2 + z^2 - r^2$ $(r > 0)$;
(ii) $\alpha x^2 + \beta y^2 - 1$ $(\alpha, \beta > 0)$;
(iii) $\alpha x^2 + \beta y^2 + \gamma z^2 - 1$ $(\alpha, \beta, \gamma > 0)$;
(iv) $\alpha x^2 + \beta y^2 - \gamma z^2 - 1$ $(\alpha, \beta, \gamma > 0)$;
(v) $\alpha x^2 + \beta y^2 - \gamma z^2 + 1$ $(\alpha, \beta, \gamma > 0)$;
(vi) $\alpha x^2 + \beta y^2 - \gamma z^2$ $(\alpha, \beta, \gamma > 0)$;
(vii) $\alpha x^2 - \beta y^2 - z$ $(\alpha, \beta > 0)$;
(viii) $xy - z$;
(ix) $\alpha x^2 + \beta y^2 - z$ $(\alpha, \beta > 0)$;
(x) $\alpha x^2 - \beta y^2 - 1$ $(\alpha, \beta > 0)$;
(xi) $\alpha x^2 - z$ $(\alpha > 0)$;
(xii) $\left((x^2 + y^2)^{1/2} - a\right)^2 + z^2 - r^2$ $(a > r > 0)$.

Sauf pour (vi) et (xii), on pourra prendre $\Omega = \mathbb{R}^3$; pour (vi) et (xii), $\Omega = \mathbb{R}^3 \setminus \{0\}$ convient. Les noms de ces surfaces sont : (i) sphère, (ii) cylindre elliptique, (iii) ellipsoïde, (iv) hyperboloïde à une nappe, (v) hyperboloïde à deux nappes, (vi) cône, (vii) et (viii) paraboloïde hyperbolique, (ix) paraboloïde elliptique, (x) cylindre hyperbolique, (xi) cylindre parabolique, (xii) tore. Toutes ces surfaces, sauf (v), (vi) et (x), sont connexes. Pour (vi), si l'on prenait $\Omega = \mathbb{R}^3$, 0 ne serait plus une valeur régulière de f correspondante ; le cône avec son sommet 0 n'est plus une surface régulière. Le tore (xii) est la surface engendrée par la rotation du cercle de centre $(0, a, 0)$ et de rayon $r < a$ dans le plan (y, z) autour de l'axe des z ; vérifier que l'on obtient la même surface en considérant $f^{-1}(0)$ où f est la fonction polynomiale, définie sur $\Omega = \mathbb{R}^3$, suivante :

$$\left(x^2 + y^2 + z^2 + a^2 - r^2\right)^2 - 4a^2(x^2 + y^2) ;$$

pour cette fonction f, 0 est une valeur régulière. (Toutes ces surfaces reviendrons dans la suite.)

[2] Soit $f(x, y, z) = x^2 + y^2 - z^2$, $(x, y, z) \in \mathbb{R}^3$. Pour tout c, $f^{-1}(c)$ est non vide ; c est une valeur régulière si $c \neq 0$; $c = 0$ est une valeur singulière ; mais tous les points de $f^{-1}(0)$, sauf $(x, y, z) = (0, 0, 0)$, sont des points réguliers de f.

[3] Soit $f \in C^1(\Omega)$, Ω ouvert non vide dans $\mathbb{R}^N$; soit S l'ensemble de points singuliers de f dans Ω. Démontrer que S est fermé dans Ω. Avec $f(x) = \sin 1/x$, $\Omega =]0, 1[\subset \mathbb{R}$, voir que S n'est pas nécessairement fermé.

4 Faire l'exercice 3 avec $f \in C^1(\Omega; \mathbb{R}^M)$.

5 Soit $f(x,y,z) = (x^2 - y^2, xy, xz, yz)$, $(x,y,z) \in \mathbb{R}^3$. Vérifier que $f : \mathbb{R}^3 \to \mathbb{R}^4$ est tel que

(i) $\operatorname{rang} f'(x,y,z) = 3$ si $xyz \neq 0$;
(ii) $\operatorname{rang} f'(x,y,z) = 2$ si $xyz = 0$, $x^2 + y^2 + z^2 > 0$;
(iii) $f'(0,0,0) = 0$.

Soit $S^2 = \{(x,y,z) : x^2 + y^2 + z^2 = 1\} \subset \mathbb{R}^3$. Le sous-ensemble $f(S^2) \subset \mathbb{R}^4$ est une surface régulière compacte dans $\mathbb{R}^4$ que l'on identifie avec $\mathbb{P}^2(\mathbb{R})$, le ***plan projectif réel***. En effet, vérifier que : $f(\xi) = f(-\xi)$ si $\xi \in S^2$; si $\eta \in f(S^2)$, $f^{-1}(\eta) \cap S^2 = \{\xi, -\xi\}$ pour un $\xi \in S^2$; $f|_{S^2} \to \mathbb{R}^4$ est une application de rang 2 dans le sens suivant :

$$f \circ \varphi : D \to \mathbb{R}^4\,,\ \varphi(x,y) = \left(x, y, \sqrt{1 - x^2 - y^2}\right),\ D = \left\{(x,y) : x^2 + y^2 < 1\right\},$$

est de rang 2 (φ est une paramétrisation de l'hémisphère nord de S^2) et de même pour les autres points de S^2. (Géométriquement, $\mathbb{P}^2(\mathbb{R})$ est S^2 où l'on a « identifié » ξ et $-\xi$; la discussion de cet exercice montre que $f(S^2)$ est obtenue par un plongement de $\mathbb{P}^2(\mathbb{R})$ dans $\mathbb{R}^4$. Il faut noter que $\mathbb{P}^2(\mathbb{R})$ ne peut pas être plongée dans $\mathbb{R}^3$, c'est-à-dire que c'est une surface que l'on ne peut pas construire dans $\mathbb{R}^3$.)

6 Soit $f(x,y,z) = x^3+y^3+z^3-3xyz$, $(x,y,z) \in \mathbb{R}^3$. Vérifier que $\nabla f(x,y,z) = (0,0,0)$ si et seulement si $x = y = z$; en conclure que tout nombre réel $c \neq 0$ est une valeur régulière de f et que $f^{-1}(c)$ est une surface régulière dans $\mathbb{R}^3$ si $c \neq 0$. Par contre 0 est une valeur singulière de f et $f^{-1}(0)$ est la réunion du plan d'équation $x + y + z = 0$ (une surface régulière) et de la droite d'équation $x = y = z$ (une courbe régulière). (Noter que $f = gh$, $g = x + y + z$, $h = x^2 + y^2 + z^2 - xy - xz - yz$; $h = 0 \Leftrightarrow x = y = z$.)

7 Soit $f = (f_1, f_2)$ où $f_1(x,y,z) = x^2 + 2xz - 2yz - y^2$ et $f_2(x,y,z) = 2x - y + z$, $(x,y,z) \in \mathbb{R}^3$. Vérifier que $\operatorname{rang} f'(x,y,z) = 2$ si $x - 2y - z \neq 0$ et $\operatorname{rang} f'(x,y,z) = 1$ si $x - 2y - z = 0$. Le point $(0,b) \in \mathbb{R}^2$, $b \neq 0$, est une valeur régulière de f, et $(0,0) \in \mathbb{R}^2$ est une valeur singulière de f. L'ensemble $f^{-1}(0,b)$, $b \neq 0$, est la réunion de deux droites parallèles (donc, une sous-variété différentielle de dimension 1) mais $f^{-1}(0,0)$ est la réunion de deux droites s'intersectant au point $(0,0,0)$ (donc, n'est pas une sous-variété différentielle). (Noter que $f_1(x,y,z) = (x - y)(x + y + 2z)$.)

8 Soient

$$f_1(x,y) = \frac{x^2 - y^2}{x^2 + y^2}\,,\quad f_2(x,y) = \frac{2xy}{x^2 + y^2}\,,\quad (x,y) \in \mathbb{R}^2 \setminus (0,0) = \Omega\,;$$

posons $f = (f_1, f_2)$. Vérifier que $\operatorname{rang} f'(x,y) = 1$, $(x,y) \in \Omega$ de sorte que f_1, f_2 sont fonctionnellement dépendantes. Déterminer $\varphi : \mathbb{R}^2 \to \mathbb{R}$ de classe C^1,

non identiquement nulle, telle que $\varphi \circ f = 0$. (Voir que $f(\Omega)$ est une courbe régulière dans $\mathbb{R}^2$ que vous pouvez identifier.)

9 Soient f_1, f_2, f_3 trois fonctions polynomiales en $(x, y, z) \in \mathbb{R}^3$ données par $(x + y - z)$, $(x - y + z)$, $(x^2 + y^2 + z^2 - 2yz)$ respectivement ; posons $f = (f_1, f_2, f_3)$. Vérifier que $\operatorname{rang} f'(x, y, z) = 2$ pour tout (x, y, z), de sorte que f_1, f_2, f_3 sont fonctionnellement dépendantes. Déterminer $\varphi : \mathbb{R}^3 \to \mathbb{R}$ de classe C^1, non identiquement nulle, telle que $\varphi \circ f = 0$. (Voir que $f(\mathbb{R}^3)$ est une surface bien connue.)

10 Soit $f : \Omega \to \mathbb{R}$ une fonction de classe C^1 où Ω est ouvert dans $\mathbb{R}^n$; si $\nabla f(x) \neq 0$, $x \in \Omega$, démontrer qu'il existe $f_2, \ldots, f_n$ dans $C^1(\Omega)$ telles que $f_1 = f, f_2, \ldots, f_n$ forment des coordonnées locales pour Ω. (Prendre $f_2(x) = x_2, \ldots, f_n(x) = x_n$ si $x = (x_1, \ldots, x_n)$.)

11 Généraliser l'exercice 10 en partant de $(f_1, \ldots, f_p)$, $p < n$.

4.5 EXTREMUMS LIÉS

4.5.1 Introduction

Un problème qui se présente fréquemment en pratique est le suivant : déterminer $\sup f(x)$ ou $\inf f(x)$ lorsque la variable x varie dans une partie (ouverte) Ω (de $\mathbb{R}^n$) assujettie à certaines conditions du type $g_1(x) = 0, \ldots, g_m(x) = 0$. Ce problème est généralement traité sous la rubrique ***extremums liés***. Un problème plus général est celui de déterminer $\sup f(x)$ ou $\inf f(x)$ lorsque la variable x est astreinte à certaines conditions supplémentaires du type $h_1(x) \leqslant 0, \ldots, h_s(x) \leqslant 0$; les problèmes de ce genre sont appelés problèmes de ***programmation*** (ou ***optimisation***) ***non linéaires*** ; la dénomination ***programmation linéaire*** concerne le cas particulier où f, g_i, h_j sont des fonctions linéaires. La difficulté pratique des problèmes de programmation, linéaires ou non linéaires, réside dans le fait que ces problèmes concernent un nombre de variables, d'équations et d'inégalités très grand. Nous considérons d'abord le problème d'extremum lié.

Nous faisons les hypothèses suivantes : $f : \Omega \to \mathbb{R}$ une fonction dérivable, Ω ouvert non vide dans $\mathbb{R}^n$, $g_i : \Omega \to \mathbb{R}$ de classe C^1, $i = 1, \ldots, m$ avec $m < n$; nous supposons, de plus, que $g = (g_1, \ldots, g_m)$ est une *submersion* de Ω dans $\mathbb{R}^m$, c'est-à-dire $\operatorname{rang} g'(x) \equiv m$; ceci veut dire que le système des fonctions $g_1, \ldots, g_m$ est fonctionnellement indépendant ; aussi, $g^{-1}(0)$ est alors une sous-variété différentielle E de dimension $n - m$ contenue dans Ω. Le problème est d'étudier $\sup\{f(x) : x \in E\}$ (le cas de inf étant parfaitement similaire : on passe d'un problème à l'autre en passant de f à $-f$). En fait, on commence par l'étude du maximum ou minimum local de f sur E. Pour illustrer ces notions précisément, disons que $a \in E$ est un point de maximum relatif (ou local) sur E pour f s'il existe un voisinage ouvert U de a tel que $f(x) \leqslant f(a)$ pour

$x \in U \cap E$; s'il y a inégalité stricte, on dira que a est un point de maximum local strict. Comme dans le cas d'extemum sans condition, on étudie d'abord tous les points de maximum relatif sur E pour passer ensuite à l'étude de $\sup f$ sur E. Dans la discussion théorique nous nous occupons seulement du problème de la détermination des points d'extremum relatifs sur E ; l'étude de $\sup f$ sur E est ensuite achevée par des considérations supplémentaires, généralement par celles de compacité comme on l'a vu dans le chapitre 2 (sect. 2.4).

Une condition nécessaire pour que le point $a \in E$ soit un point d'extremum relatif sur E pour f est obtenue par la considération suivante : soit $\gamma :]-\varepsilon, \varepsilon[\to E$ une fonction dérivable telle que $\gamma(0) = a$; la fonction $f \circ \gamma$ aura un extremum relatif au point $t = 0$ si a est un point d'extremum relatif pour f sur E ; intuitivement, c'est l'étude de f sur le chemin, passant par a, tracé par γ. Analytiquement, on doit avoir

$$g_1\big(\gamma(t)\big) = 0, \ldots, g_m\big(\gamma(t)\big) = 0, \quad -\varepsilon < t < \varepsilon, \tag{4.22}$$

pour que γ soit dans E et on doit avoir

$$\frac{\mathrm{d}}{\mathrm{d}t}(f \circ \gamma)(t)\Big|_{t=0} = f'\big(\gamma(0)\big)\gamma'(0) = \nabla f(a) \cdot \gamma'(0) = 0 \tag{4.23}$$

pour que $t = 0$ soit un point critique de $f \circ \gamma$.

De (4.22) on a, par dérivation,

$$\nabla g_i(a) \cdot \gamma'(0) = 0, \quad i = 1, \ldots, m,$$

ce qui veut dire que $\xi = \gamma'(0)$ est orthogonal (dans $\mathbb{R}^n$) au sous-espace vectoriel S (dans $\mathbb{R}^n$) engendré par les m vecteurs linéairement indépendants $\nabla g_i(a)$ $i = 1, \ldots, m$ (se rappeler que les m lignes de la matrice canonique de $g'(a)$ sont $\nabla g_i(a)$ $i = 1, \ldots, m$ et que $\operatorname{rang} g'(a) = m$). En fait, un vecteur ξ est de la forme $\gamma'(0)$ si et seulement si $\xi \in S^{\perp}$ (ξ orthogonal à S) ; nous discuterons ce point ci-après. Ainsi, (4.23) peut être interprété comme suit :

$$\nabla f(a) \in \big(S^{\perp}\big)^{\perp} = S.$$

Ceci veut dire que $\nabla f(a)$ est une combinaison linéaire des $\nabla g_i(a)$: il existe $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_m$ réels tels que

$$\nabla f(a) = \widetilde{\lambda}_1 \nabla g_1(a) + \cdots + \widetilde{\lambda}_m \nabla g_m(a). \tag{4.24}$$

Donc, *une condition nécessaire pour que $a \in E$ soit un point d'extremum local dans E pour f est que (4.24) soit vérifiée* (g étant une submersion : $\operatorname{rang} g'(x) = m$).

Il reste à voir que tout $\xi \in S^{\perp}$ est de la forme $\gamma'(0)$ avec $\gamma :]-\varepsilon, \varepsilon[\to E$, $\gamma(0) = a$, γ dérivable. Puisque E est une variété différentielle de dimension $n - m$ de classe C^1 (dans $\Omega \subset \mathbb{R}^n$), il existe une partie ouverte $W \subset \mathbb{R}^{n-m}$, un voisinage ouvert V de $a \in E$ et une paramétrisation de classe C^1, $\psi : W \to V \cap E$; on peut supposer que $0 \in W$, que W est un voisinage rectangulaire de $0 \in \mathbb{R}^{n-m}$ et que $\psi(0) = a$.

Soit $\delta_j :]-\varepsilon, \varepsilon[\to E$, $(j = 1, \ldots, n-m)$ le chemin défini par

$$t_j \longmapsto \psi(0, \ldots, 0, t_j, 0, \ldots, 0) ;$$

posons $\eta_j = \delta_j'(0) = \mathrm{D}_j\psi(0)$; puisque $g_i \circ \psi = 0$, $i = 1, \ldots, m$, on a que $g_i(\delta_j(t)) \equiv 0$ d'où

$$\nabla g_i(a) \cdot \eta_j = 0\,, \quad j = 1, \ldots, n-m\,,\ i = 1, \ldots, m\,,$$

de sorte que η_j, $j = 1, \ldots, n-m$ sont dans $S^\perp$. Puisque la matrice canonique de $\psi'(0)$ est formée des $n-m$ colonnes $\mathrm{D}_j\psi(0) = \delta_j'(0)$, $j = 1, \ldots, n-m$, et $\operatorname{rang}\psi'(0) = n-m$, on en déduit que η_j, $j = 1, \ldots, n-m$, sont linéairement indépendants, donc forment une base pour $S^\perp$ car la dimension de $S^\perp$ est $n-m$. Si $\xi \in S^\perp$, on peut alors écrire $\xi = \sum_j c_j\eta_j$, $c = (c_1, \ldots, c_{n-m})$; le chemin δ défini par

$$\delta(t) = \psi(tc)\,, \quad -\varepsilon < t < \varepsilon$$

pour $\varepsilon > 0$ suffisamment petit sera tel que $\delta(0) = a$ avec $\delta :]-\varepsilon, \varepsilon[\to E$, et

$$\delta'(0) = \psi'(0)\,c = \sum_{j=1}^{n-m} c_j\eta_j = \xi\,. \qquad \blacklozenge$$

La condition nécessaire (4.24) est traditionnellement formulée d'une certaine manière, très commode pour les applications, que l'on appelle la méthode des ***multiplicateurs de Lagrange***. On la présente comme suit.

4.5.2 Méthode des multiplicateurs de Lagrange

Les notations étant celles qui précèdent, posons

$$F(x;\lambda) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x)$$

pour $x \in \Omega$, $\lambda = (\lambda_1, \ldots, \lambda_m)$. *Une condition nécessaire pour que $a \in \Omega$ soit un extremum local dans $E = g^{-1}(0)$ pour f est qu'il existe $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_m$ tels que pour $x = a$ et pour $\widetilde{\lambda} = (\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_m)$ on a*

$$\frac{\partial F}{\partial x_j} = 0\,, \quad j = 1, \ldots, n \tag{4.25}$$

$$\frac{\partial F}{\partial \lambda_i} = 0\,, \quad i = 1, \ldots, m\,. \tag{4.26}$$

En effet, (4.26) dit que

$$g_i(a) = 0\,, \quad i = 1, \ldots, m$$

ce qui équivaut à $a \in E$; (4.25) se traduit en

$$\mathrm{D}_j f(a) - \sum_{i=1}^{m} \widetilde{\lambda}_i \, \mathrm{D}_j g_i(a) = 0 \,, \quad j = 1, \ldots, n \,,$$

ce qui est équivalent à (4.24). ♦

Si f et les g_i sont de classe C^2, on peut développer des conditions suffisantes pour qu'un point $a \in E$ remplissant (4.24) soit un extremum relatif ; il suffit d'étudier la matrice hessienne de la fonction réelle $f \circ \psi$ où ψ est une paramétrisation de E autour de a comme ci-dessus. Nous n'entrerons pas dans cette discussion.

4.5.3 Optimisation non linéaire

La position du problème est la suivante : soient $f, g_1, \ldots, g_m, h_1, \ldots, h_s$ des fonctions réelles de classe C^1 définies dans une partie ouverte non vide Ω de $\mathbb{R}^n$; nous supposons que $m < n$; à trouver des conditions nécessaires pour que $a \in \Omega$ soit un extremum local pour f dans E où

$$E = \left\{ x \in \Omega : g_i(x) = 0 \,,\ i = 1, \ldots, m \,;\ h_j(x) \leqslant 0 \,,\ j = 1, \ldots, s \right\}$$

est supposé être non vide. Pour fixer les idées, nous considérons le cas d'un minimum local.

Supposons que U soit un voisinage ouvert de a tel que $U \subset \Omega$ et

$$a \in E \cap U \quad \text{et} \quad f(a) \leqslant f(x) \,,\ x \in E \cap U \,;$$

supposons que

$$h_1(a) = \ldots = h_p(a) = 0 \,, \quad h_{p+1}(a) < 0, \ldots, h_s(a) < 0$$

($0 \leqslant p \leqslant s$; $p = 0$ signifie que $h_j(a) < 0$ pour tout j ; $p = s$ signifie que $h_j(a) = 0$ pour tout j).

Alors il existe $\lambda_0, \lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_p$ *nombres réels qui ne sont pas tous nuls tels que :*

(i) pour $j = 1, \ldots, n$

$$\lambda_0 \mathrm{D}_j f(a) + \sum_{i=1}^{m} \lambda_i \mathrm{D}_j g_i(a) + \sum_{r=1}^{p} \mu_r \mathrm{D}_j h_r(a) = 0 \,; \tag{4.27}$$

(ii) $\lambda_0 \geqslant 0$, $\mu_r \geqslant 0$, $r = 1, \ldots, p$;

(iii) si l'ensemble des vecteurs

$$\left\{ \nabla g_i(a) \,,\ i = 1, \ldots, m \,;\ \nabla h_r \,,\ r = 1, \ldots, p \right\} \subset \mathbb{R}^n$$

est linéairement indépendant (de sorte que $m + p \leqslant n$*), alors on peut choisir* $\lambda_0 = 1$ *dans (i).*

Notons que, en absence des fonctions h_r, la condition (iii) est simplement la condition que $\operatorname{rang} g'(a) = m$, $g = (g_1, \ldots, g_m)$ et l'énoncé précèdent est équivalent à la condition nécessaire pour un extremum lié donnée dans (4.24) (sauf qu'ici nous faisons l'hypothèse plus forte que f est de classe C^1).

Pour la démonstration, nous supposons, sans perte de généralité, que $a = 0$ et $f(a) = 0$. Choisissons un $\varepsilon_1 > 0$ tel que

$$\overline{B}(0;\varepsilon_1) \subset U\,, \quad h_{p+1}(x) < 0, \ldots, h_s(x) < 0 \quad \text{si } x \in \overline{B}(0;\varepsilon_1)\,.$$

Nous affirmons les deux choses suivantes :

(I) pour tout $\varepsilon \in \,]\,0\,,\,\varepsilon_1\,]$, il existe un nombre réel N tel que

$$f(x) + \|x\|^2 + N\left\{\sum_{i=1}^{m} g_i^2(x) + \sum_{r=1}^{p}\big(h_r^+(x)\big)^2\right\} > 0 \quad \text{si } \|x\| = \varepsilon$$

où $\varphi^+(x) = \max\big(\varphi(x), 0\big)$ pour toute fonction réelle φ ;

(II) pour tout $\varepsilon \in \,]\,0\,,\,\varepsilon_1\,]$, il existe un $b \in \mathbb{R}^n$, $\|b\| < \varepsilon$, et des nombres réels $\lambda_0, \lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_p$ tels que $\lambda_0 \geqslant 0$, $\mu_1 \geqslant 0, \ldots, \mu_p \geqslant 0$ avec

$$\sum_{i=0}^{m} \lambda_i^2 + \sum_{r=1}^{p} \mu_r^2 = 1$$

et tels que, si $b = (b_1, \ldots, b_n)$, on a

$$\lambda_0\big\{\mathrm{D}_j f(b) + 2b_j\big\} + \sum_{i=1}^{m} \lambda_i\, \mathrm{D}_j g_i(b) + \sum_{r=1}^{p} \mu_r\, \mathrm{D}_j h_r(b) = 0\,, \quad 1 \leqslant j \leqslant m\,.$$

En admettant (II), on obtient le résultat énoncé directement. En effet, soit $\varepsilon_1 > \varepsilon_2 > \ldots \to 0$; pour chaque ε_q, prenons un $b(q)$ et des nombres $\lambda_i(q)$, $\mu_r(q)$, comme postulés dans (II). En passant à une sous-suite appropriée, si nécessaire, nous pouvons admettre que $\lambda_i(q) \to \lambda_i$, $\mu_r(q) \to \mu_r$ lorsque $q \to \infty$ pour $i = 0, 1, \ldots, m$, $r = 1, \ldots, p$; évidemment, $b(q) \to 0$ si $q \to \infty$ car $\big\|b(q)\big\| < \varepsilon_q$; puisque les λ_i, μ_r sont tels que leur somme des carrés est 1, ces nombres ne sont pas tous zéro ; de plus, $\lambda_0 \geqslant 0$, $\mu_1 \geqslant 0, \ldots, \mu_p \geqslant 0$. Avec ces $\lambda_0, \lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_p$ et $a = 0$, on voit que (i) et (ii) de l'énoncé sont vérifiés ; sous la condition d'indépendance linéaire de (iii), λ_0 sera > 0 car autrement (i) contredirait l'indépendance linéaire des vecteurs concernés par (iii).

Etablissons maintenant (I) en raisonnant par l'absurde. Si (I) n'est pas vrai, on aurait une suite de nombres réels $N_1, N_2, \ldots \to \infty$ et une suite de vecteurs $x_1, x_2, \ldots$ dans $\mathbb{R}^n$ avec $\big\|x_j\big\| = \varepsilon$, $j \geqslant 1$, tels que pour tout j,

$$f(x_j) + \|x_j\|^2 \leqslant -N_j\left\{\sum_{i=1}^{m} g_i^2(x_j) + \sum_{r=1}^{p}\big(h_r^+(x_j)\big)^2\right\}. \tag{4.28}$$

En passant à une sous-suite appropriée, on peut supposer que $x_j \to \xi$ lorsque $j \to \infty$; alors $\|\xi\| = \varepsilon$ et $f(x_j) \to f(\xi)$ lorsque $j \to \infty$. En divisant les deux côtés dans (4.28) par N_j et en laissant ensuite $j \to \infty$, on obtiendrait que

$$\sum_{i=1}^{m} g_i^2(\xi) + \sum_{r=1}^{p} \big(h_r^+(\xi)\big)^2 = 0$$

ce qui veut dire $g_i(\xi) = 0$, $i = 1, \ldots, m$ et $h_r(\xi) \leqslant 0$, $r = 1, \ldots, p$; ainsi, $\xi \in E \cap U$ et

$$\lim_{j\to\infty} f(x_j) = f(\xi) \geqslant f(0)$$

ce qui contredit (4.28) selon laquelle $f(x_j) \leqslant -\varepsilon^2$, $j \geqslant 1$. Ceci établit (I).

Passons à l'affirmation (II). Avec N comme dans l'affirmation (I), définissons $F(x)$, $x \in U$, comme suit :

$$F(x) = f(x) + \|x\|^2 + N\left\{\sum_{i=1}^{m} g_i^2(x) + \sum_{r=1}^{p} \big(h_r^+(x)\big)^2\right\}.$$

Notons que $F(0) = 0$. Soit $b \in \overline{B}(0;\varepsilon)$ tel que

$$F(b) = \inf\big\{F(x) : x \in \overline{B}(0;\varepsilon)\big\} ;$$

alors $F(b) \leqslant F(0) = 0$ de sorte que, selon l'affirmation (I), $|b| < \varepsilon$. On en conclut que

$$\mathrm{D}_j F(b) = 0\,, \quad j = 1, \ldots, n\,.$$

On vérifie que

$$\mathrm{D}_j\big(h_r^+(x)\big)^2 = 2h_r^+(x)\,\mathrm{D}_j h_r(x) \tag{4.29}$$

en considérant séparément les cas $h_r(x) > 0$, $h_r(x) < 0$ et $h_r(x) = 0$. Alors, pour $j = 1, \ldots, n$, on aura

$$\mathrm{D}_j f(b) + 2b_j + 2N\left\{\sum_{i=1}^{m} g_i(b)\,\mathrm{D}_j g_i(b) + \sum_{r=1}^{p} h_r^+(b)\,\mathrm{D}_j h_r^+(b)\right\} = 0\,. \tag{4.30}$$

Posons

$$L = \left\{1 + \sum_{i=1}^{m} \big(2Ng_i(b)\big)^2 + \sum_{r=1}^{p} \big(2Nh_r^+(b)\big)^2\right\}^{1/2}$$

$$\lambda_0 = \frac{1}{L}\,, \quad \lambda_i = \frac{2Ng_i(b)}{L}\,,\ 1 \leqslant i \leqslant m\,,\,, \quad \mu_r = \frac{2Nh_r^+(b)}{L}\,,\ 1 \leqslant r \leqslant p\,.$$

Grâce à (4.29), $\mu_r \geqslant 0$, $1 \leqslant r \leqslant p$; en vue de (4.30), l'affirmation (II) est établie.

Ceci achève la démonstration de l'énoncé. Nous avons suivi l'exposé de E. J. McShane (« The Lagrange multiplier rule », *American Mathematical Monthly*, vol. 80, pp. 922-925, 1973). Cette démonstration donne une preuve indépendante et directe pour la validité de la méthode des multiplicateurs de Lagrange en évitant toute considération des fonctions implicites.

Dans la littérature du sujet les conditions (i), (ii) s'appellent les ***conditions de Kuhn-Tucker*** et la condition (iii) (qui permet le choix $\lambda_0 = 1$ dans (i)) est souvent appelée la ***contrainte de Kuhn-Tucker***. Les démonstrations usuelles utilisent les propriétés de convexité dans $\mathbb{R}^n$; elles restent utiles et importantes pour l'étude de ce genre de problèmes.

Notons que si les fonctions f et h_r sont toutes convexes (définies dans Ω convexe et ouvert dans $\mathbb{R}^n$) en plus d'être de classe C^1 et si les $g_i \equiv 0$, alors *les conditions (i)-(iii) sont aussi des conditions suffisantes* pour que le point a soit un minimum global de f dans E. En effet, si $x \in E$, on aura, sous conditions (i)-(iii) (où $\lambda_0 = 1$ et $\lambda_1, \ldots, \lambda_m$ sont absents),

$$
\begin{aligned}
f(x) &\geqslant f(x) + \sum_{r=1}^{p} \mu_r h_r(x) \quad (\text{car } h_r(x) \leqslant 0,\ \mu_r > 0) \\
&\geqslant \{f(a) + \nabla f(a) \cdot (x-a)\} + \sum_{r=1}^{p} \mu_r \{h_r(a) + \nabla h_r(a) \cdot (x-a)\} \\
&\qquad\qquad (\text{par la convexité des } f,\ h_r\ ;\ \text{prop. } 3.7.2) \\
&= f(a) + \sum_{r=1}^{p} \mu_r h_r(a) \quad (\text{selon condition (i)}) \\
&= f(a) \quad (\text{car } h_r(a) = 0,\ 1 \leqslant r \leqslant p)\,.
\end{aligned}
$$

Faisons une remarque simple sur les conditions (i) et (ii). Il est évident que si les conditions sont de type $h_r(x) \geqslant 0$ on aura alors $\mu_r \leqslant 0$, $1 \leqslant r \leqslant p$ dans (i) et (ii) (passer aux fonctions $-h_r$) ; de même, pour que a soit un point de maximum local (en maintenant les conditions $h_r(x) \leqslant 0$), on doit prendre $\mu_r \leqslant 0$, $1 \leqslant r \leqslant p$ (en passant au minimum local de $-f$).

Finalement, donnons une remarque sur l'ensemble E. Il se peut que celui-ci soit un ensemble discret et même fini ; ceci ne change rien dans la validité des conditions (i)-(iii). En général, on aura $m+p < n$ de sorte que sous la condition (iii), E sera une variété différentielle de dimension $n-(m+p)$ localement autour du point a.

4.5.4 Exercices

1 Trouver les extremums des fonctions f suivantes sous les conditions indiquées ; décider s'il s'agit d'un maximum ou minimum, local ou global.

(a) $f(x,y) = x^\alpha + y^\alpha$, $x + y = 1$, $x > 0$, $y > 0$, $\alpha > 0$, $\alpha \neq 1$.
(b) $f(x,y) = xy$, $x^\alpha + y^\alpha = 1$, $x > 0$, $y > 0$, $\alpha > 0$.

(c) $f(x,y,z) = ax^2 + by^2 + cz^2$, $x+y+z = 1$, $a > 0$, $b > 0$, $c > 0$.
(d) $f(x,y,z) = xyz$, $x+y+z = a$, $xy+xz+yz = b$ avec $a = 5$, $b = 8$.

Pour (d), voir que $a^2 - 2b > 0$ est une condition nécessaire et suffisante pour que les deux conditions imposées donnent un ensemble non réduit à un point (qui est, en fait, un cercle dans $\mathbb{R}^3$).

(e) $f(x,y,z) = xyz$, $x+y+z = 0$.
(f) $f(x,y,z) = (s-x)(s-y)(s-z)$, $x+y+z = 2s$, x, y, z, $s > 0$ (s fixe) ; en déduire que, parmi les triangles de périmètre fixe $2s$, c'est un triangle équilatère qui rend l'aire maximale.

2 Soient $(a,b) \in \mathbb{R}^2$, $f(x,y) = (x-a)^2+(y-b)^2$, $C = \{(x,y) : x^2+y^2 = 1\}$. Calculer $\inf\{f(x,y) : (x,y) \in C\}$; interpréter votre résultat géométriquement.

3 Soient $a \in \mathbb{R}^n$, $f(x) = \|x-a\|^2$, $S = \varphi^{-1}(0)$ où $\varphi : \mathbb{R}^n \to \mathbb{R}$ est de classe C^1 ayant 0 comme une valeur régulière telle que $\varphi^{-1}(0) \neq \emptyset$; supposons que $a \notin S$. Démontrer que si $f(\alpha) = \inf\{f(x) : x \in S\}$, $\alpha \in S$, alors $\nabla\varphi(x) = t(\alpha-a)$ pour un $t \in \mathbb{R}^*$. Interpréter votre résultat géométriquement. Déterminer α et $f(\alpha)$ si $\varphi(x) = \|x\|^2 - 1$.

4 Soient $f(x,y) = \alpha x + \beta y$, $(\alpha,\beta) \neq (0,0)$, $(x,y) \in \mathbb{R}^2$; posons $\mu = \inf\{f(x,y) : x^2+y^2 \leqslant 1\}$, $M = \sup\{f(x,y) : x^2+y^2 \leqslant 1\}$. Démontrer que μ ou M ne peut être atteint que pour (x,y) tel que $x^2+y^2 = 1$; calculer μ, M en utilisant la méthode des multiplicateurs de Lagrange. Interpréter votre résultat géométriquement et par rapport à l'inégalité de Schwarz.

5 Généraliser l'exercice 4 à $\mathbb{R}^n$ en prenant $f(x) = \alpha \cdot x$, $\alpha \neq 0$, α, x dans $\mathbb{R}^n$ et, en posant $\mu = \inf\{f(x) : \|x\| \leqslant 1\}$, $M = \sup\{f(x) : \|x\| \leqslant 1\}$.

6 Soient $x = (x_1, \ldots, x_n)$, $f(x) = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, $g(x) = \sum_i \alpha_i x_i$ avec $\alpha_i > 0$, $i = 1, \ldots, n$, $\sum_i \alpha_i = 1$, $\Omega = \{x \in \mathbb{R}^n : x_1 > 0, \ldots, x_n > 0\}$. Vérifier que l'unique extremum local dans Ω pour f sous la condition $g = 1$ se trouve au point $x_1 = \ldots = x_n = 1$. En déduire que $\sup\{f(x) : x \in \Omega,\ g(x) = 1\} = 1$ et $(1, \ldots, 1) = \xi$ est l'unique point de Ω où $g(\xi) = 1$, $f(\xi) = 1$. Conclure que $x_1^{\alpha_1} \cdots x_n^{\alpha_n} \leqslant \sum_i \alpha_i x_i$ pour $x_1 > 0, \ldots, x_n > 0$ avec égalité seulement si $x_1 = \ldots = x_n$ (*inégalité entre les moyennes arithmétiques et géométriques pondérées*).

7 Démontrer que

$$\sup\left\{\prod_{k=1}^n x_k : x_k \geqslant 0,\ k = 1, \ldots, n,\ \sum_{k=1}^n \alpha_k^2 x_k^2 = D^2\right\} = \frac{D^n}{n^{n/2}\prod_{k=1}^n \alpha_k}$$

où $\alpha_k > 0$, $k = 1, \ldots, n$, $D > 0$.

8 Posons

$$G(x_1, \ldots, x_n) = (x_1 \cdots x_n)^{1/n}, \quad H(x_1, \ldots, x_n) = \left(\frac{1}{n}\sum_i \frac{1}{x_i}\right)^{-1}$$

pour $x_1 > 0, \ldots, x_n > 0$; démontrer que $H(x_1, \ldots, x_n) \leqslant G(x_1, \ldots, x_n)$ avec égalité seulement si $x_1 = \ldots = x_n$ en établissant que le seul point d'extremum relatif pour $f(x) = x_1 \cdots x_n$, $x = (x_1, \ldots, x_n)$, dans $\Omega = \{x : x_1 > 0, \ldots, x_n > 0\}$ sous la condition $\sum_i x_i^{-1} = 1$ est le point $(n, \ldots, n)$ et en montrant que ce point donne $\inf\{f(x) : x \in \Omega, \sum_i x_i^{-1} = 1\}$.

($H(x)$ s'appelle la *moyenne harmonique*, $G(x)$ la *moyenne géométrique* et $A(x) = \sum_i x_i/n$ la *moyenne arithmétique* des nombres positifs $x_1, \ldots, x_n$; on a donc $H(x) < G(x) < A(x)$ sauf si $x_1 = \ldots = x_n$; dans ce cas les trois moyennes sont égales.)

9 Soit A une matrice symétrique réelle du type (n, n). Par un raisonnement utilisant la compacité, démontrer d'abord qu'il existe un $\xi \in \mathbb{R}^n$, $\|\xi\| = 1$, tel que $\sup\{Ax \cdot x : \|x\| = 1\} = A\xi \cdot \xi$. Avec la méthode des multiplicateurs de Lagrange, démontrer que si $\lambda = A\xi \cdot \xi$ alors $A\xi = \lambda\xi$. En conclure que ξ est un vecteur propre correspondant à la valeur propre maximale λ de A et $Ax \cdot x \leqslant \lambda\|x\|$, $x \in \mathbb{R}^n$. En considérant inf à la place de sup on démontre que $Ax \cdot x \geqslant \mu\|x\|$, $x \in \mathbb{R}^n$, où μ est la valeur propre minimale de A.

10 Soient A, B deux matrices symétriques réelles du type (n, n) ; supposons que B est du type strictement positif. En considérant $\sup\{Ax \cdot x : Bx \cdot x = 1\}$ et en procédant comme dans l'exercice 9, démontrer l'existence d'un $\xi \in \mathbb{R}^n$, $B\xi \cdot \xi = 1$, tel que, si $\lambda = A\xi \cdot \xi$, on a $A\xi = \lambda B\xi$ et $Ax \cdot x \leqslant \lambda Bx \cdot x$, $x \in \mathbb{R}^n$.

11 Démontrer que, si $p, q, r > 0$, $p^{-1} + q^{-1} = r^{-1}$, alors

$$\sup\left\{\sum_{i=1}^n (x_i y_i)^r : \sum_{i=1}^n x_i^p = 1, \sum_{i=1}^n y_i^q = 1, x_i > 0, y_i > 0, i = 1, \ldots, n\right\} = 1$$

en utilisant la méthode des multiplicateurs de Lagrange, le sup étant atteint seulement pour $y_i^q = x_i^p$, $i = 1, \ldots, n$. En déduire l'inégalité de Hölder dans la forme

$$\left\{\sum_{i=1}^n (x_i y_i)^r\right\}^{1/r} \leqslant \left\{\sum_{i=1}^n x_i^p\right\}^{1/p} \cdot \left\{\sum_{i=1}^n y_i^q\right\}^{1/q}$$

pour les x_i, y_i positifs, avec égalité si et seulement si $x_i^p = t y_i^q$, $i = 1, \ldots, n$, pour un $t > 0$.

12 Soit $f(X) = \det X$ où $X = [x_{ij}]$ est une matrice réelle de type (n, n), considérée comme un point de $\mathbb{R}^{n^2}$; soit $\Omega = \{X \in \mathbb{R}^{n^2} : f(X) \neq 0\}$. Vérifier que les extremums relatifs de $f(X)$, $X \in \Omega$, sous les conditions $\sum_j x_{ij}^2 = 1$, $i = 1, \ldots, n$, sont les matrices orthogonales. En déduire l'inégalité de Hadamard

$$|\det X| \leqslant \prod_{i=1}^{n} \left(\sum_{j=1}^{n} x_{ij}^2 \right)^{1/2} .$$

(Noter que $\mathrm{d}f/\mathrm{d}x_{ij}(X) = X_{ij}$ où X_{ij} est le cofacteur de x_{ij} et $\sum_j x_{ij} X_{\alpha j} = f(X)$ ou 0 selon que $i = \alpha$ ou $i \neq \alpha$.)

13 Soit $f(x, y) = x^2 + y^2$, $(x, y) \in \mathbb{R}^2$. Le minimum de f sous la condition $g = 0$ où $g(x, y) = y^2 - (x - 1)^3$ est atteint à l'unique point $(1, 0)$ mais les équations de multiplicateurs de Lagrange ne possèdent aucune solution.

4.6 REMARQUES

On a invoqué au paragraphe 4.2.9 le problème d'inversion globale. On pourrait de même, en relation avec la section 4.3, se poser la question de trouver le domaine maximal de l'existence d'une fonction implicite ; l'étude de cette question nous amène immédiatement aux questions topologiques importantes liées à la théorie des « revêtements ». Aussi, les problèmes de « bifurcation » des solutions de $F(x; u) = 0$ aux points $(a; b)$ où le théorème des fonctions implicites ne s'applique pas sont d'une grande importance, surtout dans le cas où x, u sont des variables de dimension infinie comme des fonctions.

Les théorèmes de ce chapitre sont établis sous l'hypothèses que les fonctions sont au moins de la classe C^1. On peut généraliser ce genre d'hypothèse, mais il faut savoir que sans l'hypothèse C^1 on peut avoir des applications assez bizarre. Ainsi, il existe des fonctions $f :]-1, 1[\to \mathbb{R}$ dérivables, $f'(0) \neq 0$, sans que f soit injective dans un voisinage ouvert de 0 (par exemple, $f(x) = x + 2x^2 \sin 1/x$ si $x \neq 0$, $f(0) = 0$) ; il y a des exemples « algébriques » de ce type avec $f : \mathbb{R}^2 \to \mathbb{R}^2$ (Dieudonné [13, tome 1, Ex. 2, chap. X, sect. 2]).

L'étude des sous-variétés différentielles de $\mathbb{R}^N$ est un sujet central des mathématiques. On y reviendra dans la suite. Il y a même un domaine des mathématiques entièrement consacré à ce sujet que l'on appelle « topologie différentielle » ; il se situe entre l'analyse, la géométrie différentielle et la topologie, et il est un des domaines le plus dynamique des mathématiques modernes. La théorie des « catastrophes » est une théorie très proche ; elle étudie les changements qualitatifs des ensembles $\{x : f(x, u) = 0\}$ lorsque u est variable, $x \in \mathbb{R}^N$, $u \in \mathbb{R}^M$, $f = (f_1, \ldots, f_k)$.

Les multiplicateurs de Lagrange apparaissent dans un article de Lagrange de 1760-61 : « Essai d'une nouvelle méthode pour déterminer les maxima et les minima des formules intégrales indéfinies » (Œuvres de Lagrange, vol. 1, p. 357). Cet article de Lagrange est, en fait, consacré aux problèmes dans un domaine que l'on appelle aujourd'hui « calcul des variations » ; Lagrange a exposé sa méthode longuement dans notre contexte dans son traité « Théorie des fonctions », 1813 (Œuvres de Lagrange, vol. 9-10, p. 291).

Le développement de l'optimisation non linéaire est récent. Sous l'impulsion des problèmes pratiques provenant de l'économie et de la recherche opérationnelle, cette théorie débute pendant la seconde guerre mondiale (1939-1945) et les premiers articles apparaissent à partir de 1947 ; l'article de Kuhn-Tucker est de 1951 (*Second Berkeley Symposium*). Pour quelques indications sur la variété des problèmes posés consulter Dantzig et Eaves [11] ; pour un exposé introductif à la programmation linéaire et ses applications voir de Werra [42].

Nous avons établi le théorème d'inversion locale par un raisonnement de minimisation (preuve du lemme 4.2.4) ; dans une forme plus élaborée, c'est aussi la base de la démonstration de McShane que nous avons suivie dans le paraghaphe 4.5.3. Une démonstration très répandue aujourd'hui est celle utilisant le théorème du point fixe du paragraphe 1.9.5 (par exemple Dieudonné [13, tome 1, p. 275], Spivak [39, p. 35]). Une autre démonstration (d'abord détaillée par Dini, cf. § 4.3.2) est celle consistant à résoudre les équations $F_1(x;u) = 0, \ldots, F_m(x;u) = 0$ par éliminations successives des variables $u_1, \ldots, u_m$ (cf. Douchet et Zwahlen [16, vol. 2, p. 91]).

CHAPITRE 5

Intégration

Une théorie de l'intégration des fonctions réelles définies dans $\mathbb{R}^N$ est indispensable pour tous les cours d'analyse. Après avoir rappelée la problématique sous-jacente, nous développons, dans la longue section 5.2, la théorie de l'intégrale de Lebesgue dans un espace mesuré quelconque.

L'utilité de cette théorie sera visible dans les considérations de la section 5.3 qui permettent l'interversion des limites et des dérivations sous le signe de l'intégration, et surtout dans la section 5.4 qui donne des règles extrêmement simples de l'échange de l'ordre de deux intégrations successives (théorème de Fubini), un procédé indispensable pour nombreux calculs. Les discussions des sections 5.2 à 5.4 sont parfaitement générales et elles concernent des espaces mesurés quelconques ; dans la section 5.5 et dans les sections suivantes, nous considérons le cas particulier de l'espace $\mathbb{R}^N$ muni de sa mesure naturelle, entre autres, et nous établissons en détail dans les sections 5.6 et 5.7 la formule de changement de variables dans $\mathbb{R}^N$. Les autres sections de ce chapitre donnent une multitude d'exemples d'applications de la théorie développée ; elles concernent presque exclusivement les calculs dans $\mathbb{R}^N$. Dans la section 5.11, nous indiquons par des propositions précises que les fonctions très générales dans $\mathbb{R}^N$ peuvent être approchées, de manière satisfaisante, par les fonctions de classe C^∞. La dernière section 5.12 donne, entre autres, une formule asymptotique de Stirling.

Un développement détaillé de la théorie de l'intégrale de Lebesgue est assez laborieux ; présenter toutes les démonstrations, surtout dans les sections 5.2, 5.4 et 5.5, aurait doublé le volume du présent chapitre, déjà long. Néanmoins, la présentation est suffisamment détaillée pour qu'un lecteur moyen arrive à appliquer les propositions fondamentales de la théorie aisément ; il suffit pour cela d'accepter un petit nombre d'affirmations sur la notion de mesurabilité comme exactes, d'admettre que la mesure de Lebesgue dans $\mathbb{R}^N$ existe et d'accepter la proposition 5.2.5 ; presque toutes les autres propositions et affirmations de ce chapitre sont démontrées complètement et illustrées généralement par un grand nombre d'exemples concrets. Finalement, disons clairement que la théorie est présentée dans un cadre très général, non pas dans le but d'atteindre un maximum de généralité, mais pour avoir une théorie *simple*, *exacte* et *applicable* aux problèmes de l'intégration dans $\mathbb{R}^N$.

5.1 REMARQUES PRÉLIMINAIRES

Dans cette section, nous rapellons la notion de l'intégration familière à tout lecteur de ce livre pour expliquer ensuite la nécessité d'une théorie plus générale, appelée la théorie de l'***intégrale de Lebesgue***. Cette théorie n'est pas indispensable pour les considérations de notre livre, mais elle rend un grand nombre de raisonnements et de calculs plus simples à justifier.

5.1.1 L'intégrale de Riemann : rappels intuitifs

Soit f une fonction réelle, définie dans un intervalle J de $\mathbb{R}$. Un problème fondamental est le suivant : trouver une fonction, $F : J \to \mathbb{R}$, telle que $F'(x) = f(x)$, $x \in J$. Une telle fonction F s'appelle une ***primitive*** de f. Il est bien connu (et facile à démontrer à partir du théorème des accroissements finis) que deux primitives de f ne diffèrent entre elles que par une constante additive. La question théorique importante est de savoir si une primitive F *existe*. Bien entendu, pour certaines fonctions simples (comme $f(x) = x^n$ ou $\sin x$ ou $\cos x$) on peut obtenir une primitive immédiatement en étudiant les formules élémentaires de dérivation. Ensuite, en se servant des règles générales de dérivation, on peut obtenir explicitement les primitives des fonctions plus compliquées comme

$$f(x) = \frac{x}{(x^2+1)^2}, \quad f(x) = \mathrm{e}^{2x}\cos x, \ \ldots$$

Ces règles de dérivation transposées au calcul des primitives s'appellent généralement : ***intégration par substitution***, ***intégration par parties***, etc. Mais aucune de ces manipulations ne peut démontrer que *chaque fonction continue* f possède une primitive F. En fait, on sait bien que la plupart des fonctions « élémentaires », même très simples, ne possèdent pas de primitives « élémentaires » (considérer, par exemple, $f(x) = \exp x^2$). Ce qu'il nous faut, c'est obtenir l'*existence* d'une primitive F, étant donnée une fonction continue quelconque $f : J \to \mathbb{R}$. Cette existence s'établit en introduisant l'***intégrale définie*** de Cauchy (en cas de f continue) et de Riemann (au cas de f plus général). Si $a < b$ sont deux points de J, on définit d'abord une ***somme de Riemann*** de f, $S(f;P)$, associée à une ***partition de Riemann*** P de $[a, b]$ comme suit :

$$P = \{x_0 = a, x_1, \ldots, x_n = b\}, \quad x_0 < x_1 < \ldots < x_n,$$

$$S(f;P) = \sum_{i=1}^{n} \alpha_i \lambda(A_i)$$

où $A_i = [x_{i-1}, x_i]$, $\lambda(A_i) = x_i - x_{i-1}$, $\alpha_i = f(\xi_i)$, ξ_i étant un point dans A_i. Pour chaque partition de Riemann, il y a une infinité de sommes de Riemann

associées, correspondant aux différents choix des points ξ_i dans A_i. Posons $\|P\| = \max_i \lambda(A_i)$; on dit que

$$\lim_{\|P\|\to 0} S(f;P)$$

existe et est égale à $c \in \mathbb{R}$, si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que (pour tout choix des $\xi_i \in A_i$) :

$$\left|S(f;P) - c\right| < \varepsilon$$

dès que $\|P\| < \delta$. On démontre que le nombre c est uniquement défini, si la limite existe, et on pose

$$\int_a^b f = \int_a^b f(x)\,\mathrm{d}x = c = \lim_{\|P\|\to 0} S(f;P)\,;$$

le nombre $\int_a^b f$ s'appelle l'***intégrale définie*** de f sur $[a\,,\,b]$ et on dit que f est ***intégrable au sens de Riemann*** sur $[a\,,\,b]$.

Un théorème fondamental, déjà essentiellement démontré par Cauchy, est que si f est continue dans $[a\,,\,b]$ alors $\int_a^b f$ existe. Ce théorème donne une solution satisfaisante au problème de l'existence d'une primitive F pour une *fonction continue* $f : J \to \mathbb{R}$ (J un sous-intervalle de $\mathbb{R}$). En effet, on pose

$$F(x) = \int_a^x f, \quad x \in J\,,$$

en fixant un point quelconque $a \in J$ et en convenant que

$$\int_a^x f = -\int_x^a f \quad \text{si } x < a$$

et $F(a) = 0$. On démontre facilement que $F'(x) = f(x)$, $x \in J$; c'est une partie du ***théorème fondamental du calcul différentiel et intégral*** (que nous abrégerons souvent comme ***théorème fondamental***) ; l'autre partie de ce théorème dit que si $f = \varphi'$ et f est intégrable sur $[a\,,\,b]$ au sens de Riemann, alors

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b \varphi'(x)\,\mathrm{d}x = \varphi(x)\Big|_a^b = \varphi(b) - \varphi(a)\,,$$

ce qui offre un outil commode pour un calcul explicite de beaucoup d'intégrales définies. On utilise aussi le terme d'***intégrale indéfinie***, pour $\int_a^x f$, le mot « indéfini » se rapportant au fait qu'une des limites de l'intégrale est une variable x ; on a pris aussi l'habitude d'écrire $\int f(x)\,\mathrm{d}x$ pour n'importe quelle primitive de f.

On sait que $\int_a^b f$ peut bien exister (au sens de Riemann) sans que f soit continue ; nous supposons connues les propriétés élémentaires de cette intégrale.

Rappelons une interprétation géométrique importante de l'intégrale définie : si $f \geqslant 0$, $\int_a^b f$ représente l'aire de l'ensemble

$$E = \{(x, y) : a \leqslant x \leqslant b,\ 0 \leqslant y \leqslant f(x)\}.$$

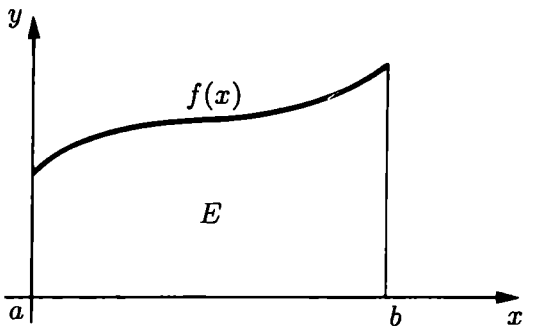


Fig. 5.1

Les débutants sont tentés de *définir* $\int_a^b f$ comme l'aire de l'ensemble E ; il faut comprendre que cette définition n'a de sens que si l'on a préalablement défini la notion d'aire d'une classe de sous-ensembles de $\mathbb{R}^2$. Il est, en fait, possible de définir précisément aire(A) pour une grande classe de sous-ensembles A de $\mathbb{R}^2$ et ensuite définir $\int_a^b f$ comme aire(E) si $f \geqslant 0$ et si E est un ensemble pour lequel aire(E) est définie ; mais ce procédé aussi demande des raisonnements longs. D'autre part, nous verrons dans la suite que la notion de l'intégrale définie permet de définir et de calculer l'aire de beaucoup de sous-ensembles de $\mathbb{R}^2$.

L'interprétation intuitive de $\int_a^b f$ comme la limite des sommes $\sum_i \alpha_i \lambda_i(A_i)$ est d'une importance capitale pour beaucoup d'applications. Toutes les intégrales que nous rencontrerons dans la suite peuvent être interprétées dans ce sens.

Un dernier mot sur le problème de la détermination d'une primitive. On a vu que si $f : J \to \mathbb{R}$ est continue (J un sous-intervalle de $\mathbb{R}$), alors f possède une primitive ; que peut-on dire pour f discontinue ? Curieusement, cette question élémentaire ne semble pas avoir une réponse simple. Si f possède une discontinuité (dite de première espèce) à un point a ($f(a_+) \neq f(a_-)$), alors f ne peut pas avoir une primitive dans un intervalle ouvert contenant a, essentiellement parce qu'une fonction dérivée g, même si elle n'est pas continue, est continue dans le sens de Darboux, ce qui veut dire qu'elle prend toutes les valeurs intermédiaires entre deux de ces valeurs prises : si $\alpha < \beta$, alors pour tout nombre c entre $g(\alpha)$ et $g(\beta)$ il existe un γ, $\alpha < \gamma < \beta$, tel que $g(\gamma) = c$. Une caractérisation de f possédant une primitive semble être difficile à obtenir. Notons d'autre part que même si f est intégrable au sens de Riemann dans $[a, b]$, la fonction F définie par

$$F(x) = \int_a^x f, \quad a \leqslant x \leqslant b,$$

n'est pas nécessairement dérivable. C'est un théorème important de la théorie de l'intégration de Lebesgue que $F'(x)$ existe « presque partout » x (dans un sens que l'on précisera dans la suite) et $F'(x) = f(x)$ presque partout dans $[a, b]$, le théorème étant valable même pour la classe très étendue des fonctions f intégrables au sens de Lebesgue.

5.1.2 Intégrales multiples : rappels intuitifs

Du point de vue pratique, le calcul d'une intégrale multiple se fait par le calcul successif des intégrales définies d'une variable. Par exemple,

$$\iint_{\substack{a<x<b \\ c<y<d}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \left(\int_c^d f(x,y)\,\mathrm{d}y \right) \mathrm{d}x = \int_c^d \left(\int_a^b f(x,y)\,\mathrm{d}x \right) \mathrm{d}y \tag{5.1}$$

où a, b, c, d sont des constantes, finies ou infinies.

La formule (5.1) est valable sous des conditions très générales ; il suffit que $f \geqslant 0$ ou que f soit bornée et a, b, c, d soient finies et que f soit une fonction *mesurable*, une notion que nous expliquerons dans la suite. Aucune condition de continuité sur f n'est nécessaire pour (5.1) ; il suffit de travailler avec une définition adéquate de l'intégrale, celle que nous donnerons dans la suite. Disons, pour l'instant, que toutes les fonctions que l'on rencontrera en pratique seront mesurables ; c'est même une tâche difficile de démontrer qu'il existe des fonctions non mesurables.

La formule (5.1) fait partie d'un théorème important que l'on appelle ***théorème de Fubini***. Pour une fonction de n variables, on a une formule analogue affirmant (sous des conditions très générales) l'égalité entre les $n!$ intégrales itérées.

5.1.3 Nécessité d'une théorie de l'intégration générale

Pour que le théorème de Fubini (5.1) soit valable dans la généralité indiquée (ce qui est très commode dans d'innombrables calculs), il est essentiel d'avoir une théorie générale comme celle de l'intégrale de Lebesgue. C'est une des raisons pour l'introduction dans la suite de ce genre d'intégration. La restriction à une intégrale de Riemann compliquerait notre travail avec des formules du type (5.1).

Une autre raison pour l'adoption de l'intégrale de Lebesgue est qu'elle facilite la justification des formules du type suivant :

$$\lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b \lim_{n\to\infty} f_n(x)\,\mathrm{d}x \tag{5.2}$$

$$\sum_{n=1}^{\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b \sum_{n=1}^{\infty} f_n(x)\,\mathrm{d}x \tag{5.3}$$

(sous des conditions simples et naturelles).

Nous nous référerons à (5.2) pour parler de la formule de l'***échange de limite et intégrale***, et nous désignerons (5.3) comme la formule de l'***intégration (d'une série) terme à terme***.

Il ne faut pas penser que (5.2) et (5.3) soient toujours valables ; il y a en fait de nombreux contre-exemples simples que l'on rencontrera dans la suite ; ce qu'il faut comprendre c'est que, dans la théorie de Lebesgue, il y a des conditions simples qui garantissent la validité de (5.2) et (5.3).

Pour donner un exemple élémentaire où (5.2) est évidemment valable (selon le théorème de convergence dominée de Lebesgue donné dans la suite) mais dont la démonstration est difficile dans le cadre de la théorie de Riemann, considérons l'énoncé suivant : soit $\{f_n\}$ une suite de fonctions réelles, continues, définies dans l'intervalle $[0, 1]$ telle que $0 \leqslant f_n(x) \leqslant 1$, $\lim_n f_n(x) = 0$, $x \in [0, 1]$; alors

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x = 0\,.$$

On notera que l'énoncé précédent est géométriquement presque évident ; de plus, il est immédiatement vérifiable selon le théorème de Lebesgue cité ; mais une démonstration directe utilisant la définition de l'intégrale de Riemann est assez laborieuse.

5.2 INTÉGRALE DE LEBESGUE

Pour éviter les définitions séparées de l'intégrale d'une fonction d'une variable et celle d'une fonction de plusieurs variables, nous procéderons par la construction d'une théorie de l'intégration pour les fonctions $f : X \to \mathbb{R}$ où X est un ensemble quelconque muni d'une mesure positive μ (cf. § 5.2.1, 5.2.2). Ici aussi, l'intégrale de f par rapport à μ, notée

$$\int f\,\mathrm{d}\mu \quad \text{ou} \quad \int_X f\,\mathrm{d}\mu \quad \text{ou} \quad \int_X f(x)\mu(\mathrm{d}x)\,,$$

sera essentiellement une limite des expressions

$$\sum_{i=1}^{n} \alpha_i\,\mu(A_i)\,.$$

Dans les applications considérées dans ce livre, X sera toujours une partie d'un $\mathbb{R}^N$. Mais dans la théorie des probabilités, on a besoin d'un espace d'épreuves X qui n'est pas toujours un sous-ensemble d'un $\mathbb{R}^N$ de manière naturelle.

Pour l'exposé de ce livre, le développement d'une théorie dans un ensemble abstrait X est motivé par la commodité de la présentation ; en même temps, cette théorie générale n'est pas plus difficile à développer que si nous nous restreignons à $\mathbb{R}^N$ ou même à $\mathbb{R}$.

5.2.1 Espaces boréliens

Soit X un ensemble non vide quelconque. Une famille Σ de parties de X s'appelle une ***tribu*** si les conditions suivantes sont satisfaites :

- $\emptyset \in \Sigma$, $X \in \Sigma$;
- si $A \in \Sigma$ alors $A^c \in \Sigma$;
- si $A_n \in \Sigma$, $n = 1, 2, \ldots$, alors $\bigcup_{n=1}^{\infty} A_n \in \Sigma$.

On vérifie immédiatement que si $A_n \in \Sigma$, $n = 1, 2, \ldots$, alors $\bigcap_n A_n \in \Sigma$ car $A_n^c \in \Sigma$ et

$$\left(\bigcap_n A_n\right)^c = \bigcup_n A_n^c \in \Sigma .$$

Un ensemble X muni d'une tribu Σ de ses parties s'appelle un ***espace borélien*** (aussi, ***espace mesurable***) ; voici quelques exemples d'une tribu dans X :

- $\Sigma = \{\emptyset, X\}$ (appelée ***tribu triviale*** dans X) ;
- $\Sigma =$ la famille de toutes les parties de X (notée 2^X) ;
- $\Sigma = \left\{\bigcup_{n\in I} A_n\,,\ I \in 2^{\mathbb{N}}\right\}$ où $A_0, A_1, \ldots$ est une partition de X, c'est-à-dire $X = \bigcup_{n\in\mathbb{N}} A_n$ et les A_i sont deux à deux disjoints.

Si $\mathcal{F}$ est une famille quelconque de parties de X, il existe une plus petite tribu Σ contenant toutes les parties de $\mathcal{F}$ dans le sens que

$$\Sigma \text{ est une tribu, } \Sigma \supset \mathcal{F} \text{ ; si } \Sigma' \text{ est une tribu, } \Sigma' \supset \mathcal{F} \text{ alors } \Sigma' \supset \Sigma .$$

On dit que Σ est la tribu ***engendrée*** par $\mathcal{F}$. L'existence de Σ est démontrée en observant que

$$\Sigma = \bigcap \{\Sigma' : \Sigma' \text{ une tribu}\,,\ \Sigma' \supset \mathcal{F}\}\,,$$

car l'intersection d'une famille de tribu est aussi une tribu ; noter que 2^X est une tribu contenant $\mathcal{F}$. Soit $X \subset \mathbb{R}^N$; la ***tribu borélienne*** de X est la tribu engendrée par les parties relativement ouvertes de X. Rappeler que $A \subset X$ est relativement ouvert dans X, si $A = X \cap U$ où U est une partie ouverte de $\mathbb{R}^N$.

Soit Σ_X la tribu borélienne de $X \subset \mathbb{R}^N$; *si Σ est la tribu borélienne de $\mathbb{R}^N$, alors* $\Sigma_X = \Sigma \cap X$ dans le sens que $A \in \Sigma_X$ si et seulement si $A = B \cap X$ où B est un ***ensemble borélien*** dans $\mathbb{R}^N$, c'est-à-dire $B \in \Sigma$. Les parties de Σ_X s'appellent les ***ensembles boréliens*** de X.

La tribu borélienne de $\mathbb{R}^N$ contient (presque) tous les ensembles d'un intérêt analytique. Par définition, une partie ouverte de $\mathbb{R}^N$ est borélienne ; par complémentation, on voit que tous les ensembles fermés sont boréliens ; donc les réunions et les intersections dénombrables formées des parties ouvertes ou fermées sont aussi des ensembles boréliens.

Il est difficile de construire ou donner un exemple concret d'un sous-ensemble de $\mathbb{R}^N$ qui ne soit pas borélien. On sait qu'il existe des ensembles boréliens $A \subset \mathbb{R}^2$ tels que $\pi_1(A)$ n'est pas borélien dans $\mathbb{R}^1$ où $\pi_1 : \mathbb{R}^2 \to \mathbb{R}^1$ est l'application projection sur l'axe x_1 : $\pi_1(x_1, x_2) = x_1$. La situation est analogue pour les projections dans $\mathbb{R}^N$; plus généralement, il existe des fonctions continues $f : \mathbb{R}^N \to \mathbb{R}^M$ tels que $f(A)$ n'est pas nécessairement borélien dans $\mathbb{R}^M$ même si A est borélien dans $\mathbb{R}^N$.

Néanmoins, dans ce livre, on va *travailler comme si* tous les ensembles que nous rencontrons étaient boréliens.

5.2.2 Mesures

Soit (X, Σ) un espace borélien. Une ***mesure positive*** (dans X ou sur Σ) est une application $\mu : \Sigma \to [0, \infty]$ satisfaisant les conditions suivantes :

- $\mu(\emptyset) = 0$;
- si $A_n \in \Sigma$, $n = 1, 2, \ldots$, est une suite d'ensembles deux à deux disjoints alors

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

(cette propriété de μ s'appelle la ***σ-additivité*** ou l'***additivité dénombrable*** de μ).

La mesure μ s'appelle ***finie*** si $\mu(X) < \infty$; si μ est une mesure finie, alors $\mu(A) < \infty$ pour tout $A \in \Sigma$, car toute mesure positive est ***monotone*** dans le sens que si $A \subset B$, A, B dans Σ, alors $\mu(A) \leqslant \mu(B)$ (cf. Ex. 1(i), § 5.2.15). Si $\mu(X) = 1$, on dit que μ est une ***mesure de probabilité*** ; dans ce cas, $\mu(A) \leqslant 1$ pour $A \in \Sigma$.

Une mesure positive μ sur Σ s'appelle ***σ-finie*** si l'espace entier X peut être écrit comme

$$X = A_1 \cup A_2 \cup \cdots, \quad A_n \in \Sigma, \ n \geqslant 1,$$

avec $\mu(A_n) < \infty$, $n \geqslant 1$. Cette condition de σ-finitude de μ peut être reformulée comme suit :

$$X = B_1 \cup B_2 \cup \cdots, \quad B_n \in \Sigma, \ \mu(B_n) < \infty, \ n \geqslant 1,$$

où les B_n sont deux à deux disjoints. En effet, il suffit de poser, dans la définition qui précède,

$$B_1 = A_1, \ B_2 = A_2 \setminus A_1, \ B_3 = A_3 \setminus (A_1 \cup A_2), \ \ldots$$

pour obtenir les B_n deux à deux disjoints nécessaires ; remarquer que, pour tout $n \geqslant 1$,

$$B_1 \cup \cdots \cup B_n = A_1 \cup \cdots \cup A_n.$$

Ce passage d'une réunion d'une suite d'ensembles quelconque à une réunion « équivalente » d'une suite d'ensembles deux à deux disjoints est souvent utile.

La plupart des mesures positives en analyse sont des mesures σ-finies ou finies. On parle aussi des mesures réelles ou vectorielles (qui sont des applications $\mu : \Sigma \to \mathbb{R}$ ou $\mathbb{R}^n$ ou un autre espace vectoriel ayant la propriété de σ-additivité et telle que $\mu(\emptyset) = 0$) mais dans ce livre nous n'en ferons aucun usage.

Un espace borélien (X, Σ) muni d'une mesure positive μ s'appelle un ***espace mesuré*** (X, Σ, μ) (***fini*** si $\mu(X) < \infty$, ***σ-fini*** si μ est σ-fini).

Etant donné un espace mesuré (X, Σ, μ), chaque $A \in \Sigma$ détermine un espace mesuré (que l'on pourrait appeler un sous-espace mesuré de X mais cette terminologie n'est pas courante et elle ne sera pas utilisée) selon les définitions suivantes : on pose

$$\Sigma_A = \{B \subset A : B \in \Sigma\}, \quad \mu_A(B) = \mu(B) \text{ si } B \in \Sigma_A .$$

On voit immédiatement que $B \in \Sigma_A$ si et seulement si B est de la forme $\widetilde{B} \cap A$ où $\widetilde{B} \in \Sigma$ et μ_A est la restriction de μ aux ensembles de Σ_A. En fait, nous n'utiliserons pas cette notation, qui est peut-être logiquement nécessaire mais pratiquement assez inutile ; si nécessaire, nous dirons que nous considérons la ***restriction de μ à A*** (seulement si $A \in \Sigma$, sans introduire une nouvelle notation) : c'est une terminologie un peu illogique mais brève et sans ambiguïté.

Il est remarquable que, pour une grande partie de la théorie de l'intégration (§ 5.2.4-5.2.11), on n'ait besoin d'aucune condition sur la mesure positive μ. C'est dans le théorème de Fubini (sect. 5.4) que la condition de σ-finitude sur μ joue un rôle. Comme les seules mesures positives dont nous aurons besoin dans ce livre sont les mesures de Lebesgue dans $\mathbb{R}^N$ pour $n \geqslant 1$ et ses restrictions, la condition de σ-finitude sur μ ne nous posera aucun problème dans les applications.

L'exemple le plus simple d'une mesure positive est la ***mesure de comptage*** dans un ensemble X quelconque ; on muni X de la tribu 2^X de toutes ses parties, et on pose $\mu(A) = \infty$ si A est une partie infinie et $\mu(A) = n$ si A est une partie finie de n éléments ; on écrit aussi $\mu(A) = \text{card}(A)$. La mesure de comptage dans X est σ-finie si et seulement si X est un ensemble au plus dénombrable. On trouve d'autres exemples de ce type dans les exercices du paragraphe 5.2.15.

Il est difficile de fournir explicitement d'autres exemples de mesures positives. Essayons de décrire l'exemple le plus important d'une mesure en analyse, la ***mesure de Lebesgue*** dans $\mathbb{R}^N$ qui sera notée λ_N (ou λ, si la dimension de l'espace n'a pas besoin d'être précisée). Par définition, λ_N est d'abord définie dans la tribu borélienne de $\mathbb{R}^N$ de telle manière que si P est un pavé borné dans $\mathbb{R}^N$, $P = I_1 \times \cdots \times I_N$, I_j étant un sous-intervalle de $\mathbb{R}$ d'extrémités $a_j < b_j$, alors

$$\lambda_N(P) = \prod_{j=1}^{N} (b_j - a_j) . \tag{5.4}$$

On voit que λ_1 correspond à la ***longueur*** d'un intervalle dans $\mathbb{R}^1$, λ_2 à l'aire d'un rectangle dans $\mathbb{R}^2$, λ_3 au volume d'un parallélépipède rectangulaire dans $\mathbb{R}^3$. Aussi, $\lambda_N(A)$ s'appelle souvent simplement la ***mesure*** de $A \subset \mathbb{R}^N$ (ou le ***volume*** de $A \subset \mathbb{R}^3$, l'***aire*** de $A \subset \mathbb{R}^2$, etc.). Il n'est pas du tout évident que λ_N, défini pour un pavé borné par (5.4), puisse être prolongé à la tribu borélienne Σ de $\mathbb{R}^N$ de telle sorte que l'on ait une mesure. Même pour $\mathbb{R}^1$ la tâche est longue ; elle était accomplie par Lebesgue dans sa thèse publiée en 1902 (cf. Œuvres de Lebesgue). Aujourd'hui, on se base sur un théorème général du type suivant (***prolongement des mesures***).

*Soit $\mathcal{A}$ une famille de sous-ensembles d'un ensemble X telle que : $\emptyset \in \mathcal{A}$; si $A, B \in \mathcal{A}$ alors $A \setminus B \in \mathcal{A}$; si $A_1, \ldots, A_n \in \mathcal{A}$ alors $(A_1 \cup \cdots \cup A_n) \in \mathcal{A}$ ($\mathcal{A}$ s'appelle un **anneau** ou un **clan**). Soit $\mu : \mathcal{A} \to [0, \infty[$ tel que : $\mu(\emptyset) = 0$; si $A_n \in \mathcal{A}$, $n = 1, 2, \ldots$, sont deux à deux disjoints et $\bigcup_n A_n \in \mathcal{A}$ alors*

$$\mu\left(\bigcup_n A_n\right) = \Sigma_n A_n$$

(σ-additivité sur $\mathcal{A}$). On suppose aussi que

$$X = B_1 \cup B_2 \cup \cdots, \quad B_n \in \mathcal{A}.$$

Dans ces conditions, μ peut être prolongée uniquement en une mesure positive $\widetilde{\mu}$ définie sur Σ, la tribu engendrée par $\mathcal{A}$; $\widetilde{\mu}$ est σ-finie.

Dans $\mathbb{R}^1$ on prend comme $\mathcal{A}$ la réunion finie des intervalles de type $[a, b[$, a, b nombres finis, et on pose $\mu([a, b[) = b - a$. Il n'est pas trop difficile de vérifier que μ remplit les conditions de σ-additivité du théorème du prolongement ; la mesure $\widetilde{\mu}$ est alors λ_1, la mesure de Lebesgue dans $\mathbb{R}^1$. La théorie est complètement similaire pour la construction de λ_N dans $\mathbb{R}^N$.

5.2.3 Fonctions mesurables

Soit (X, Σ) un espace borélien. Une fonction $f : X \to \mathbb{R}$ s'appelle une ***fonction mesurable*** réelle (ou ***fonction Σ-mesurable***), si $f^{-1}(B) \in \Sigma$ pour toute partie borélienne $B \subset \mathbb{R}$. En particulier, si f est mesurable, les ensembles suivants sont dans Σ si $b \in \mathbb{R}$:

$$f^{-1}(b) = \{x : f(x) = b\}, \quad f^{-1}(]-\infty, b]) = \{x : f(x) \leqslant b\}$$
$$f^{-1}([b, \infty[) = \{x : f(x) \geqslant b\}, \quad f^{-1}(]-\infty, b[) = \{x : f(x) < b\}, \quad \text{etc.}$$

En fait, on démontre que *si $\{x : f(x) \leqslant b\} \in \Sigma$ pour tout $b \in \mathbb{R}$ alors f est mesurable.*

Les fonctions mesurables réelles définies dans X forment une classe de fonctions très vaste ayant des propriétés de stabilité remarquables. *Toute combinaison linéaire des fonctions mesurables est mesurable ; un produit des fonctions mesurables est mesurable ; le quotient de deux fonctions mesurables f, g*

est mesurable en définissant $f(x)/g(x) = 0$ *(par exemple), si* $g(x) = 0$ *; les limites, les limites supérieures et inférieures des fonctions mesurables sont aussi mesurables.* Naturellement, il faut démontrer ces affirmations, ce que nous ne ferons pas ici. *Grosso modo*, on peut dire que toute opération analytique effectuée sur une suite de fonctions mesurables aboutit à une fonction mesurable.

Une propriété remarquable des fonctions mesurables est la possibilité de les approcher par des fonctions ***simples***. Sur (X, Σ), une fonction $f : X \to \mathbb{R}$ s'appelle ***simple*** si $f(X)$ est un ensemble fini $\{a_1, \ldots, a_n\}$; *une telle fonction est mesurable si et seulement si* $f^{-1}(a_j) = A_j \in \Sigma$, $j = 1, \ldots, n$. On a alors

$$f = \sum_{j=1}^{n} a_j 1_{A_j}, \quad 1_A(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \in A^c, \end{cases}$$

où les a_j sont des nombres réels distincts et $A_j = f^{-1}(a_j)$ est une partition finie de X ; f est mesurable si et seulement si $A_j \in \Sigma$, $1 \leqslant j \leqslant n$. On a maintenant le ***théorème d'approximation*** suivant pour un espace borélien (X, Σ).

Si $f : X \to \mathbb{R}$ *est mesurable alors il existe une suite* $\{f_n\}$, $n = 1, 2, \ldots$, *de fonctions simples mesurables, telle que*

$$f(x) = \lim_{n \to \infty} f_n(x), \quad x \in X.$$

Si $f : X \to \mathbb{R}_+$, *on peut choisir les* f_n *ci-dessus telles que* $0 \leqslant f_1(x) \leqslant f_2(x) \leqslant \ldots$, $x \in X$.

La fonction 1_A, $A \subset X$, s'appelle la ***fonction indicatrice*** de A. Le théorème précédent dit alors que toute fonction mesurable f est la limite des combinaisons finies des fonctions indicatrices mesurables, la limite étant une limite croissante si $f \geqslant 0$.

Si $X \subset \mathbb{R}^N$ muni de sa tribu borélienne Σ_X, les fonctions mesurables $f : X \to \mathbb{R}$ s'appellent les ***fonctions boréliennes***. Il est important de savoir *qu'une fonction continue est borélienne* ; la preuve se base sur le fait que $f^{-1}(V)$ est ouvert dans X si V est ouvert dans $\mathbb{R}$ de sorte que $f^{-1}(V)$ est borélien si V est ouvert ; un petit raisonnement montre alors que $f^{-1}(B)$ est borélien dans X si B est borélien dans $\mathbb{R}$.

Pour certains raisonnements il est utile de considérer les fonctions $f : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$. Si (X, Σ) est un espace borélien, une telle fonction s'appelle mesurable si $f^{-1}(\infty)$, $f^{-1}(-\infty)$ et $f^{-1}(B)$ sont dans Σ pour $B \subset \mathbb{R}$, B borélien. Le théorème d'approximation reste valable pour de telles fonctions.

Soit $A \in \Sigma$ où (X, Σ) est un espace borélien. Une fonction $f : A \to \overline{\mathbb{R}}$ s'appelle ***mesurable*** (ou ***Σ-mesurable***) si son prolongement $\widetilde{f} : X \to \overline{\mathbb{R}}$, défini par $\widetilde{f}|_A = f$ et $\widetilde{f}(x) = 0$ si $x \notin A$, est mesurable. On peut démontrer facilement que $f : A \to \overline{\mathbb{R}}$ $(A \in \Sigma)$ est mesurable si et seulement si $f^{-1}(B) \in \Sigma$ dès que $B \subset \mathbb{R}$ est borélien et $f^{-1}(\infty)$, $f^{-1}(-\infty)$ sont dans Σ. On appellera $\widetilde{f}$ le ***prolongement par zéro*** de f.

Une fonction $f : A \to \overline{\mathbb{R}}$ s'appellera une ***fonction numérique*** ; si les valeurs de f sont dans $\mathbb{R}$ on dira que f est une ***fonction réelle***. Ainsi une fonction numérique positive ($\geqslant 0$) est une application à valeurs dans $\overline{\mathbb{R}}_+ = [0, \infty]$; une fonction positive aura ses valeurs dans $\mathbb{R}_+ = [0, \infty[$.

5.2.4 Intégrales des fonctions positives

Soit (X, Σ, μ) un espace mesuré. Si $f = 1_A$, $A \in \Sigma$, on définit

$$\int f \, d\mu = \mu(A) . \tag{5.5}$$

Si $f = \sum_{i=1}^n a_i 1_{A_i}$, $a_i \in \mathbb{R}_+$, $A_i \in \Sigma$, $i = 1, \ldots, n$, on définit

$$\int f \, d\mu = \sum_{i=1}^n a_i \mu(A_i) \tag{5.6}$$

avec la convention que $0 \cdot \infty = 0 = \infty \cdot 0$; $a + \infty = \infty + a = \infty$, si $0 \leqslant a \leqslant \infty$; $a \cdot \infty = \infty = \infty \cdot a$ si $0 < a \leqslant \infty$. Ceci définit $\int f \, d\mu$ pour $f \geqslant 0$, simple et mesurable.

Si $f : X \to \overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{\infty\}$ est mesurable, on pose

$$\int f \, d\mu = \sup \left\{ \int g \, d\mu : 0 \leqslant g \leqslant f , \ g \text{ simple, mesurable} \right\} . \tag{5.7}$$

Les trois formules (5.5) à (5.7) définissent $\int f \, d\mu$ pour $f \geqslant 0$ (éventuellement à valeur ∞) et mesurable ; notons que

$$0 \leqslant \int f \, d\mu \leqslant \infty .$$

On dit qu'une telle fonction f est ***intégrable*** (ou ***μ-intégrable***) si $\int f \, d\mu < \infty$; on dit que f est ***négligeable*** (ou ***μ-négligeable***) si $\int f \, d\mu = 0$.

Soit $A \in \Sigma$ et $f : A \to [0, \infty]$. Si f est mesurable on définit $\int_A f \, d\mu$ par la formule suivante :

$$\int_A f \, d\mu = \int \widetilde{f} \, d\mu$$

où $\widetilde{f} : X \to [0, \infty]$ est le prolongement par zéro de f (c'est-à-dire : $\widetilde{f}(x) = 0$ si $x \notin A$, $\widetilde{f}|_A = f$). Notons que si $f : X \to [0, \infty]$ est mesurable, alors

$$\int_A f \, d\mu = \int f 1_A \, d\mu .$$

Aussi, on écrira $\int_X f \, d\mu$ au lieu de $\int f \, d\mu$ pour souligner qu'il s'agit de la ***μ-intégrale de*** f sur X.

Nous résumons les propriétés essentielles de l'intégrale pour les fonctions numériques mesurables et positives, dans la proposition suivante.

5.2.5 Proposition

Soit (X, Σ, μ) un espace mesuré et $\mathcal{M}_+$ l'ensemble des fonctions $f : X \to [0, \infty]$ mesurables. L'application $f \mapsto \int f \, d\mu$, $f \in \mathcal{M}_+$, possède les propriétés suivantes :

(i) (positivité) $0 \leqslant \int f \, d\mu \leqslant \infty$, $f \in \mathcal{M}_+$;

(ii) (linéarité) si $f, g \in \mathcal{M}_+$, $a, b \in \mathbb{R}_+$, *alors*

$$\int (af + bg) \, d\mu = a \int f \, d\mu + b \int g \, d\mu ;$$

(iii) (théorème de convergence monotone) si $f_n \in \mathcal{M}_+$, $n = 1, 2, \ldots$, *est une suite monotone croissante* $(f_1 \leqslant f_2 \leqslant \ldots)$ *et* $f(x) = \lim_{n\to\infty} f_n(x)$, $x \in X$, *alors* $f \in \mathcal{M}_+$ *et*

$$\lim_{n\to\infty} \int f_n(x) \mu(dx) = \int f(x) \mu(dx) .$$

Nous ne démontrerons pas cette proposition ici. Notons que (i) est incorporée dans la définition même de l'intégrale et que (ii) est visiblement valable dans (5.6) d'après la définition de l'intégrale pour f simple dans $\mathcal{M}_+$. La clef de voûte de la théorie de l'intégration de Lebesgue est (iii) ; la plupart des autres propriétés importantes des intégrales se déduisent de (iii). Remarquons encore que (iii), formellement, est une conséquence de la propriété de σ-additivité de μ.

La proposition suivante est une conséquence immédiate de la proposition 5.2.5, mais nous la présentons séparément car elle contient les propriétés fondamentales de l'intégrale (pour les fonctions numériques, mesurables et positives) dans une forme directement applicable dans la suite.

5.2.6 Proposition

Soient (X, Σ, μ), $\mathcal{M}_+$ comme dans la proposition 5.2.5. Dans cet énoncé les fonctions sont dans $\mathcal{M}_+$ et les ensembles sont dans Σ.

(i) $\int f \, d\mu \leqslant \int g \, d\mu$ *si* $0 \leqslant f \leqslant g$.

(ii) $\int \sum_{n=1}^{\infty} f_n \, d\mu = \sum_{n=1}^{\infty} \int f_n \, d\mu$.

(iii) $\int_A f \, d\mu = \sum_{n=1}^{\infty} \int_{A_n} f \, d\mu$ *si* $A = \bigcup_{n=1}^{\infty} A_n$,
les A_n étant deux à deux disjoints.

(iv) si $f = \sup_n f_n$, $n = 1, 2, \ldots$,

$$\sup_n \int f_n \, d\mu \leqslant \int f \, d\mu \leqslant \sum_{n=1}^{\infty} \int f_n \, d\mu .$$

(v) Si $A = \bigcup_{n=1}^{\infty} A_n$,

$$\int_A f \, d\mu \leqslant \sum_{n=1}^{\infty} \int_{A_n} f \, d\mu .$$

(vi) Si $\{f_n\}$, $n = 1, 2, \ldots$, *est une suite décroissante et* $f(x) = \lim_{n\to\infty} f_n(x)$, $x \in X$, *et si* f_p *est* μ*-intégrable pour un indice* p, *alors* f *est* μ*-intégrable et*

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu .$$

(vii) (Lemme de Fatou)

$$\int \liminf_{n\to\infty} f_n \, d\mu \leqslant \liminf_{n\to\infty} \int f_n \, d\mu .$$

DÉMONSTRATION

(i) On a $g = f + h$, $h \in \mathcal{M}_+$. (Le fait que f est mesurable demande un petit raisonnement que nous ne donnons pas ici.) D'après la linéarité et la positivité de l'intégrale

$$\int g \, d\mu = \int f \, d\mu + \int h \, d\mu \geqslant \int f \, d\mu . \tag{5.8}$$

Notons pour la suite que si g est μ-intégrable, alors (5.8) entraîne que f et h sont μ-intégrables et

$$\int f \, d\mu = \int g \, d\mu - \int h \, d\mu . \tag{5.9}$$

(ii) C'est une conséquence immédiate de la linéarité de l'intégrale et le théorème de convergence monotone : posons

$$g_n = f_1 + \cdots + f_n , \quad g = \lim_{n\to\infty} g_n = \sum_{n=1}^{\infty} f_n ;$$

alors

$$\int \sum_{n=1}^{\infty} f_n \, d\mu = \int g \, d\mu = \lim_{n\to\infty} \int g_n \, d\mu = \lim_{n\to\infty} \sum_{i=1}^{n} \int f_i \, d\mu = \sum_{i=1}^{\infty} \int f_i \, d\mu .$$

(iii) C'est un cas particulier de (ii) car

$$f 1_A = \sum_{n=1}^{\infty} f 1_{A_n}$$

si les A_n sont deux à deux disjoints et $A = \bigcup_n A_n$.

(iv) Puisque

$$\max_{1\leqslant i\leqslant n} f_i \leqslant f_1 + \cdots + f_n$$

et

$$f = \sup_n f_n = \lim_{n\to\infty} \left(\max_{1\leqslant i\leqslant n} f_i \right) \leqslant \sum_{n=1}^{\infty} f_n ,$$

on a

$$\int f \, d\mu \leqslant \int \sum_{n=1}^{\infty} f_n \, d\mu = \sum_{n=1}^{\infty} \int f_n \, d\mu .$$

(v) C'est un cas particulier de (iv) car

$$f \, 1_A = \sup_n (f \, 1_{A_n}) .$$

(vi) On peut supposer, sans perte de généralité, que $p = 1$ de sorte que les f_n et f sont toutes μ-intégrables. Définissons $\varphi_n \in \mathcal{M}_+$ tel que

$$f_1 = \varphi_n + f_{n+1}$$

avec $\varphi_n(x) = 0$ si $f_1(x) = \infty$, $f_{n+1}(x) = \infty$ et $\varphi_n(x) = \infty$ si $f_1(x) = \infty$, $f_{n+1}(x) < \infty$. Alors

$$\varphi_1 \leqslant \varphi_2 \leqslant \ldots , \quad \lim_{n\to\infty} \varphi_n = \varphi , \quad f_1 = \varphi + f$$

et les φ_n, φ sont μ-intégrables ; de l'observation (5.9), on a

$$\int \varphi_n \, d\mu = \int f_1 \, d\mu - \int f_{n+1} \, d\mu$$

d'où, selon le théorème de convergence monotone,

$$\int \varphi \, d\mu = \lim_{n\to\infty} \int \varphi_n \, d\mu = \int f_1 \, d\mu - \lim_{n\to\infty} \int f_{n+1} \, d\mu . \tag{5.10}$$

De nouveau, de l'observation (5.9), on a

$$\int \varphi \, d\mu = \int f_1 \, d\mu - \int f \, d\mu$$

d'où, en comparant avec (5.10), on obtient

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu .$$

(vii) Rappelons que

$$\liminf_{n\to\infty} f_n = \sup_{n\geqslant 1} \left(\inf_{k\geqslant n} f_k \right) = \lim_{n\to\infty} g_n$$

(v) Si $A = \bigcup_{n=1}^{\infty} A_n$,

$$\int_A f \, d\mu \leqslant \sum_{n=1}^{\infty} \int_{A_n} f \, d\mu .$$

(vi) Si $\{f_n\}$, $n = 1, 2, \ldots$, *est une suite décroissante et* $f(x) = \lim_{n\to\infty} f_n(x)$, $x \in X$, *et si* f_p *est* μ*-intégrable pour un indice* p, *alors* f *est* μ*-intégrable et*

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu .$$

(vii) (Lemme de Fatou)

$$\int \liminf_{n\to\infty} f_n \, d\mu \leqslant \liminf_{n\to\infty} \int f_n \, d\mu .$$

DÉMONSTRATION

(i) On a $g = f + h$, $h \in \mathcal{M}_+$. (Le fait que f est mesurable demande un petit raisonnement que nous ne donnons pas ici.) D'après la linéarité et la positivité de l'intégrale

$$\int g \, d\mu = \int f \, d\mu + \int h \, d\mu \geqslant \int f \, d\mu . \tag{5.8}$$

Notons pour la suite que si g est μ-intégrable, alors (5.8) entraîne que f et h sont μ-intégrables et

$$\int f \, d\mu = \int g \, d\mu - \int h \, d\mu . \tag{5.9}$$

(ii) C'est une conséquence immédiate de la linéarité de l'intégrale et le théorème de convergence monotone : posons

$$g_n = f_1 + \cdots + f_n , \quad g = \lim_{n\to\infty} g_n = \sum_{n=1}^{\infty} f_n ;$$

alors

$$\int \sum_{n=1}^{\infty} f_n \, d\mu = \int g \, d\mu = \lim_{n\to\infty} \int g_n \, d\mu = \lim_{n\to\infty} \sum_{i=1}^{n} \int f_i \, d\mu = \sum_{i=1}^{\infty} \int f_i \, d\mu .$$

(iii) C'est un cas particulier de (ii) car

$$f 1_A = \sum_{n=1}^{\infty} f 1_{A_n}$$

si les A_n sont deux à deux disjoints et $A = \bigcup_n A_n$.

(iv) Puisque

$$\max_{1\leqslant i\leqslant n} f_i \leqslant f_1 + \cdots + f_n$$

et

$$f = \sup_n f_n = \lim_{n\to\infty} \left(\max_{1\leqslant i\leqslant n} f_i \right) \leqslant \sum_{n=1}^{\infty} f_n ,$$

on a

$$\int f \, \mathrm{d}\mu \leqslant \int \sum_{n=1}^{\infty} f_n \, \mathrm{d}\mu = \sum_{n=1}^{\infty} \int f_n \, \mathrm{d}\mu .$$

(v) C'est un cas particulier de (iv) car

$$f \, 1_A = \sup_n (f \, 1_{A_n}) .$$

(vi) On peut supposer, sans perte de généralité, que $p = 1$ de sorte que les f_n et f sont toutes μ-intégrables. Définissons $\varphi_n \in \mathcal{M}_+$ tel que

$$f_1 = \varphi_n + f_{n+1}$$

avec $\varphi_n(x) = 0$ si $f_1(x) = \infty$, $f_{n+1}(x) = \infty$ et $\varphi_n(x) = \infty$ si $f_1(x) = \infty$, $f_{n+1}(x) < \infty$. Alors

$$\varphi_1 \leqslant \varphi_2 \leqslant \ldots , \quad \lim_{n\to\infty} \varphi_n = \varphi , \quad f_1 = \varphi + f$$

et les φ_n, φ sont μ-intégrables ; de l'observation (5.9), on a

$$\int \varphi_n \, \mathrm{d}\mu = \int f_1 \, \mathrm{d}\mu - \int f_{n+1} \, \mathrm{d}\mu$$

d'où, selon le théorème de convergence monotone,

$$\int \varphi \, \mathrm{d}\mu = \lim_{n\to\infty} \int \varphi_n \, \mathrm{d}\mu = \int f_1 \, \mathrm{d}\mu - \lim_{n\to\infty} \int f_{n+1} \, \mathrm{d}\mu . \tag{5.10}$$

De nouveau, de l'observation (5.9), on a

$$\int \varphi \, \mathrm{d}\mu = \int f_1 \, \mathrm{d}\mu - \int f \, \mathrm{d}\mu$$

d'où, en comparant avec (5.10), on obtient

$$\lim_{n\to\infty} \int f_n \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu .$$

(vii) Rappelons que

$$\liminf_{n\to\infty} f_n = \sup_{n\geqslant 1} \left(\inf_{k\geqslant n} f_k \right) = \lim_{n\to\infty} g_n$$

où $g_n = \inf_{k \geqslant n} f_k$ est une suite croissante. Donc selon le théorème de convergence monotone

$$\int \liminf_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int g_n \, d\mu \leqslant \lim_{n\to\infty} \left(\inf_{k\geqslant n} \int f_k \, d\mu \right) = \liminf_{n\to\infty} \int f_n \, d\mu$$

en remarquant que $g_n \leqslant f_k$ si $k \geqslant n$ de sorte que

$$\int g_n \, d\mu \leqslant \int f_k \, d\mu \leqslant \inf_{k\geqslant n} \int f_k \, d\mu . \qquad \blacklozenge$$

REMARQUES. L'énoncé (vi) de la proposition 5.2.6 peut être en défaut si aucune des f_n n'est μ-intégrable. En effet, si $X = \mathbb{N}$ muni de sa mesure de comptage μ et $f_n = 1_{A_n}$, $A_n = \{n, n+1, \ldots\}$, $n > 1$, alors on voit que $f_n \to 0$, en décroissant, lorsque $n \to \infty$, mais

$$\int f_n \, d\mu = \mu(A_n) = \infty , \quad n \geqslant 1 .$$

L'usage du lemme de Fatou est illustré par la situation typique suivante : supposons que $f_n \to f$ si $n \to \infty$ et $\int f_n \, d\mu \leqslant M < \infty$ pour $n \geqslant 1$; alors on conclut que f est μ-intégrale car $\int f \, d\mu \leqslant M$.

5.2.7 Intégrales des fonctions non nécessairement positives

Soit (X, Σ, μ) un espace mesuré.

Si $f : X \to \overline{\mathbb{R}}$ est une fonction numérique quelconque, définissons f^+ (la ***partie positive*** de f) et f^- (la ***partie négative*** de f) comme suit :

$$f^+(x) = \max\big(f(x), 0\big) , \quad f^-(x) = \max\big(-f(x), 0\big)$$

avec pour convention

$$\max(\infty, 0) = \infty , \quad \max(-\infty, 0) = 0 , \quad -(-\infty) = \infty .$$

Notons que f^+, f^- sont des fonctions numériques positives ayant les propriétés suivantes :

$$f(x) = f^+(x) - f^-(x) , \quad \big|f(x)\big| = f^+(x) + f^-(x) ,$$

avec pour convention naturelle

$$\infty - a = \infty , \quad a - \infty = -\infty \quad \text{si } a \in \mathbb{R} .$$

Observer que si $f^+(x) > 0$ alors $f^-(x) = 0$, et si $f^-(x) > 0$ alors $f^+(x) = 0$; aussi, $f^+(x) = f^-(x) = 0$ si $f(x) = 0$ de sorte que $f^+(x) f^-(x) = 0$ pour tout $x \in X$ (se rappeler de la convention $0 \cdot \infty = 0$). De plus, si f est mesurable, f^+ et f^- sont aussi mesurables.

Une fonction mesurable $f : X \to \overline{\mathbb{R}}$ s'appelle ***μ-intégrable*** si chacune des fonctions positives f^+, f^- est μ-intégrable et on pose alors

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu . \tag{5.12}$$

De (5.11), on voit que

$$\int |f| \, d\mu = \int f^+ \, d\mu + \int f^- \, d\mu$$

pour toute fonction mesurable $f : X \to \overline{\mathbb{R}}$ et *que f est μ-intégrable si et seulement si* $\int |f| \, d\mu < \infty$. De cette observabion, on voit qu'une fonction mesurable $f : X \to \overline{\mathbb{R}}$ est μ-intégrable si et seulement s'il existe une fonctoin $\varphi : X \to \overline{\mathbb{R}}_+$ mesurable et μ-intégrable telle que $|f| \leqslant \varphi$.

On notera par $\mathcal{L}^1(\mu)$ ou $\mathcal{L}^1(X, \mu)$ ou $\mathcal{L}^1(X, \Sigma, \mu)$ la classe des fonctions *réelles* et μ-intégrables $f : X \to \mathbb{R}$.

La proposition suivante résume certaines propriétés principales de $\mathcal{L}^1(\mu)$ et de l'application $f \mapsto \int f \, d\mu$. Comme dans le cas des fonctions positives on posera

$$\int_A f \, d\mu = \int f \, 1_A \, d\mu \quad (0 \cdot -\infty = -\infty \cdot 0 = 0)$$

pour toute fonctions μ-intégrable f et $A \in \Sigma$; si $f : A \to \overline{\mathbb{R}}$, $A \in \Sigma$, on définit $\widetilde{f}$ comme le prolongement de f par 0 en dehors de A et on dira que f est ***μ-intégrable*** (***dans*** A) si $\widetilde{f}$ est μ-intégrable ; aussi, on posera

$$\int_A f \, d\mu = \int \widetilde{f} \, d\mu .$$

Si nécessaire, on écrira $\int_X f \, d\mu$ au lieu de $\int f \, d\mu$ pour indiquer que l'intégration de f a bien lieu sur l'espace X tout entier.

5.2.8 Proposition

Soit (X, Σ, μ) un espace mesuré.

(i) $\mathcal{L}^1(\mu)$ est un espace vectoriel réel et l'application $f \mapsto \int f \, d\mu$ est une application linéaire de $\mathcal{L}^1(\mu)$ dans $\mathbb{R}$; de plus,

$$\left| \int f \, d\mu \right| \leqslant \int |f| \, d\mu , \quad f \in \mathcal{L}^1(\mu) .$$

(ii) (Monotonicité) Si f, g, h sont dans $\mathcal{L}^1(\mu)$ et $g \leqslant f \leqslant h$, alors

$$\int g \, d\mu \leqslant \int f \, d\mu \leqslant \int h \, d\mu .$$

(iii) (Théorème de convergence dominée) Soit $\{f_n\}$, $n = 1, 2, \ldots$, *une suite de fonctions de* $\mathcal{L}^1(\mu)$ *telle que*

$$\lim_{n\to\infty} f_n(x) = f(x)\,,\ x \in X \quad et \quad |f_n(x)| \leqslant M(x)\,,\ x \in X\,,\ n \geqslant 1$$

où $M : X \to \overline{\mathbb{R}}_+$ *est* μ*-intégrable. Alors*

$$\lim_{n\to\infty} \int |f_n - f|\,\mathrm{d}\mu = 0\,, \quad \lim_{n\to\infty} \int f_n\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu$$

et

$$\lim_{n\to\infty} \int_A f_n\,\mathrm{d}\mu = \int_A f\,\mathrm{d}\mu \quad \textit{uniformément pour } A \in \Sigma\,.$$

DÉMONSTRATION

(i) Si f, $g \in \mathcal{L}^1(\mu)$ et α, $\beta \in \mathbb{R}$ l'inégalité

$$|\alpha f + \beta g| \leqslant |\alpha|\,|f| + |\beta|\,|g|$$

montre que $(\alpha f + \beta g) \in \mathcal{L}^1(\mu)$; donc $\mathcal{L}^1(\mu)$ est un espace vectoriel. Pour la linéarité de $f \mapsto \int f\,\mathrm{d}\mu$ il suffit de voir que

$$\int \alpha f\,\mathrm{d}\mu = \alpha \int f\,\mathrm{d}\mu\,, \quad \int (f+g)\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu + \int g\,\mathrm{d}\mu\,.$$

Posons $h = f + g$; alors

$$h = h^+ - h^- = (f^+ - f^-) + (g^+ - g^-)$$

d'où

$$h^+ + f^- + g^- = h^- + f^+ + g^+\,;$$

par la linéarité de $\int$ pour les fonctions positives, on a

$$\int h^+\,\mathrm{d}\mu + \int f^-\,\mathrm{d}\mu + \int g^-\,\mathrm{d}\mu = \int h^-\,\mathrm{d}\mu + \int f^+\,\mathrm{d}\mu + \int g^+\,\mathrm{d}\mu$$

ce qui donne

$$\begin{aligned}\int h\,\mathrm{d}\mu &= \int h^+\,\mathrm{d}\mu - \int h^-\,\mathrm{d}\mu \\ &= \left(\int f^+\,\mathrm{d}\mu - \int f^-\,\mathrm{d}\mu\right) + \left(\int g^+\,\mathrm{d}\mu - \int g^-\,\mathrm{d}\mu\right) \\ &= \int f\,\mathrm{d}\mu + \int g\,\mathrm{d}\mu\,.\end{aligned}$$

Toutes ces manipulations sont justifiées car elles ne concernent que les nombres réels.

La relation

$$\int \alpha f \, d\mu = \alpha \int f \, d\mu$$

est évidente si $\alpha = 0$; si $\alpha > 0$, elle se déduit immédiatement des formules

$$(\alpha f)^+ = \alpha \cdot f^+, \quad (\alpha f)^- = \alpha \cdot f^- ;$$

pour $\alpha = -1$, la relation est vraie par une vérification directe des définitions ; la relation est valable donc pour $\alpha < 0$ aussi car $\alpha f = |\alpha| \cdot (-f)$. Finalement, (5.12) donne

$$\left| \int f \, d\mu \right| \leqslant \int f^+ \, d\mu + \int f^- \, d\mu = \int |f| \, d\mu, \quad f \in \mathcal{L}^1(\mu)$$

grâce à (5.11).

(ii) Notons que $f - g \geqslant 0$, $h - f \geqslant 0$ et que $f - g$, $h - f$ sont μ-intégrables ; de la linéarité de l'intégrale, on aura

$$0 \leqslant \int (f - g) \, d\mu = \int f \, d\mu - \int g \, d\mu$$

d'où $\int g \, d\mu \leqslant \int f \, d\mu$; de même, on a $\int f \, d\mu \leqslant \int h \, d\mu$.

(iii) Posons $g_n = |f_n - f|$, $h_n = \sup_{k \geqslant n} g_k$. On voit que

$$|f(x)| \leqslant M(x), \quad 0 \leqslant g_n(x) \leqslant 2M(x), \; 0 \leqslant h_n(x) \leqslant 2M(x), \; x \in X$$

d'où on déduit la μ-intégrabilité de f, g_n, h_n, $n \geqslant 1$. Puisque $g_n \to 0$ lorsque $n \to \infty$ on a que $h_n \to 0$ lorsque $n \to \infty$; de plus, $\{h_n\}$ est une suite de fonctions décroissante :

$$h_1(x) \geqslant h_2(x) \geqslant \ldots, \quad x \in X .$$

De (vi) de la proposition 5.2.6 on obtient que

$$\lim_{n \to \infty} \int h_n \, d\mu = 0 ;$$

d'où, en tenant compte de $g_n \leqslant h_n$, on déduit que

$$\lim_{n \to \infty} \int |f_n - f| \, d\mu = 0 .$$

L'inégalité

$$\left| \int_A f_n \, d\mu - \int_A f \, d\mu \right| \leqslant \int_X |f_n - f| \, d\mu$$

établit le reste. ♦

Il est utile de rassembler un certain nombre de conséquences immédiates de la proposition 5.2.8 dans la forme de la proposition suivante.

5.2.9 Proposition

Soit (X, Σ, μ) un espace mesuré.

(i) Soit $\{f_n\}$, $n = 1, 2, \ldots$, une suite de fonctions de $\mathcal{L}^1(\mu)$ telle que $\sum_{n=1}^{\infty} f_n(x)$ converge pour $x \in X$; supposons qu'il existe une fonction $M : X \to \overline{\mathbb{R}}_+$, μ-intégrable, telle que

$$\left|f_1(x) + \cdots + f_n(x)\right| \leqslant M(x), \quad x \in X, \ n \geqslant 1.$$

Alors, si $f = \sum_n f_n$, on a $f \in \mathcal{L}^1(\mu)$ et

$$\int f \, \mathrm{d}\mu = \sum_{n=1}^{\infty} \int f_n \, \mathrm{d}\mu.$$

De plus,

$$\lim_{n\to\infty} \int \left| f - \sum_{i=1}^{n} f_i \right| \mathrm{d}\mu = 0$$

et

$$\lim_{n\to\infty} \int_A (f_1 + \cdots + f_n) \, \mathrm{d}\mu = \int_A f \, \mathrm{d}\mu$$

uniformément pour $A \in \Sigma$. Aussi

$$\int |f| \, \mathrm{d}\mu \leqslant \sum_{n=1}^{\infty} \int |f_n| \, \mathrm{d}\mu \leqslant \infty$$

(la série $\sum_n \int |f_n| \, \mathrm{d}\mu$ peut bien diverger).

(ii) Soit $f \in \mathcal{L}^1(\mu)$; si $A_n \in \Sigma$, $n = 1, 2, \ldots$, est une suite d'ensembles deux à deux disjoints et $A = \bigcup_{n=1}^{\infty} A_n$ alors

$$\int_A f \, \mathrm{d}\mu = \sum_{n=1}^{\infty} \int_{A_n} f \, \mathrm{d}\mu.$$

(iii) Si $\{f_n\}$, $n = 1, 2, \ldots$, est une suite de fonctions de $\mathcal{L}^1(\mu)$ qui converge uniformément vers une fonction f et si $\mu(X) < \infty$, alors $f \in \mathcal{L}^1(\mu)$ et

$$\lim_{n\to\infty} \int |f_n - f| \, \mathrm{d}\mu = 0, \quad \lim_{n\to\infty} \int f_n \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu.$$

(iv) Si $f = \sum_{n=1}^{\infty} f_n$, $f_n \in \mathcal{L}^1(\mu)$, $n = 1, 2, \ldots$, et la série converge uniformément alors, si $\mu(X) < \infty$, $f \in \mathcal{L}^1(\mu)$ et

$$\lim_{n\to\infty} \int \left| \sum_{i=1}^{n} f_i - f \right| \mathrm{d}\mu = 0, \quad \sum_{i=1}^{\infty} \int f_i \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu.$$

DÉMONSTRATION

(i) C'est une conséquence du théorème de convergence dominée appliquée sur $\varphi_n = f_1 + \cdots + f_n$.

(ii) C'est une conséquence de (i) avec $f_n = f\,1_{A_n}$.

(iii) La convergence uniforme de f_n vers f entraîne que pour tout $\varepsilon > 0$, il existe n_ε tel que $|f_n(x) - f(x)| < \varepsilon$, $x \in X$, si $n \geqslant n_\varepsilon$; ainsi

$$\int |f_n - f|\,\mathrm{d}\mu < \varepsilon\mu(X)\,, \quad n \geqslant n_\varepsilon\,,$$

ce qui donne le résultat car $\varepsilon > 0$ est arbitraire ; la μ-intégrabilité de f est une conséquence de l'inégalité

$$|f(x)| \leqslant \varepsilon + |f_{n_\varepsilon}(x)|\,, \quad x \in X\,.$$

(iv) C'est une conséquence de (iii) appliqué sur $\varphi_n = f_1 + \cdots + f_n$. ♦

REMARQUE. Les énoncés (iii) et (iv) peuvent être en défaut si $\mu(X) = \infty$ (cf. exercices, § 5.2.15).

5.2.10 Intégrale comme la limite d'une somme

Soit (X, Σ, μ) un espace mesuré.

Si $f : X \to \mathbb{R}_+$ est mesurable, on sait que f est la limite d'une suite croissante de fonctions $f_n \geqslant 0$, simples et mesurables (§ 5.2.3, théorème d'approximation). Selon le théorème de convergence monotone, $\int f_n\,\mathrm{d}\mu \to \int f\,\mathrm{d}\mu$ si $n \to \infty$; or

$$\int f_n\,\mathrm{d}\mu = \sum_{k=1}^{N_n} a_{k,n}\,\mu(A_{k,n})$$

si f_n prend les valeurs distinctes $a_{1,n}, a_{2,n}, \ldots, a_{N_n,n}$, sur les ensembles deux à deux disjoints $A_{1,n}, A_{2,n}, \ldots, A_{N_n,n}$. Ainsi $\int f\,\mathrm{d}\mu$ est la limite des sommes finies représentant $\int f_n\,\mathrm{d}\mu$.

Si $f : X \to \mathbb{R}$ est μ-intégrable, $f = f^+ - f^-$ avec f^+, f^- positives et μ-intégrables, alors de ce qui précède on voit que $\int f\,\mathrm{d}\mu$ est, de nouveau, une limite des sommes finies associées a f^+ et f^-.

Ce qui précède donne une méthode générale pour établir les propriétés des fonctions intégrables. On établit la propriété pour une fonction indicatrice ; ensuite on passe aux fonctions positives, simples, mesurables (qui sont des combinaisons linéaires finies des fonctions indicatrices) ; un passage à la limite, en se servant du théorème de convergence monotone, établit la propriété pour les fonctions positives mesurables ; finalement on passe au cas général d'une fonction f en l'écrivant $f = f^+ - f^-$.

5.2.11 Ensembles négligeables, fonctions μ-mesurables

Soit (X, Σ, μ) un espace mesuré.

Soit $N \subset X$; on dit que N est ***μ-négligeable*** s'il existe $A \in \Sigma$, $N \subset A$, tel que $\mu(A) = 0$.

Un ensemble $E \subset X$ s'appelle ***μ-mesurable*** s'il existe $A \subset E \subset B$ tels que A, $B \in \Sigma$ et $\mu(B \setminus A) = 0$.

On peut démontrer que Σ_μ, *la famille d'ensembles μ-mesurables, est une tribu contenant la tribu Σ* ; en posant $\widetilde{\mu}(E) = \mu(A)$ ci-dessus, on obtient une mesure $\widetilde{\mu}$ sur Σ_μ qui prolonge μ. On dit que $(\Sigma_\mu, \widetilde{\mu})$ est la ***complétion*** de μ. Une fonction $f : X \to \mathbb{R}$ s'appelle ***μ-mesurable*** si $f^{-1}(B) \in \Sigma_\mu$ pour toute partie borélienne $B \subset \mathbb{R}$. On peut démontrer que $f : X \to \mathbb{R}$ *est μ-mesurable si et seulement s'il existe g, h, $g \leqslant f \leqslant h$, g et h étant mesurables, telles que $(h - g)$ est μ-négligeable* (c'est-à-dire $\int (h - g)\, \mathrm{d}\mu = 0$).

On dit qu'une propriété P est valable ***μ-presque partout*** dans A si l'ensemble E où P n'est pas valable est μ-négligeable ; si $A = X$ on dit simplement que P est valable μ-presque partout ou p.p.(μ).

Par exemple, soit $f_n(x) = x^n$, $n = 1, 2, \ldots$, $x \in \mathbb{R}$. On sait que $x^n \to 0$ si $n \to \infty$ lorsque $0 \leqslant x < 1$; on peut alors dire que $f_n \to 0$ si $n \to \infty$ p.p.(λ) dans $[0\,,\,1]$, car l'ensemble exceptionnel E ici se réduit à $\{1\}$ dont la mesure de Lebesgue est zéro.

Notons que λ_N dans $\mathbb{R}^N$ est une mesure telle que *tout ensemble contenant un nombre au plus dénombrable de points est λ_N-négligeable*. En effet, l'ensemble $\{a\}$ est négligeable car, pour tout $\varepsilon > 0$,

$$\lambda_N\big(\{a\}\big) \leqslant \big(2\varepsilon\big)^N \longrightarrow 0 \quad \text{si } \varepsilon \to 0\,;$$

cette inégalité provient du fait que $\{a\} \subset P_\varepsilon$, où P_ε est le pavé compact de centre a et côté 2ε et du fait que, dans tout espace mesuré (X, Σ, μ), $A \subset B$ entraîne que $\mu(A) \leqslant \mu(B)$ si A, B sont dans Σ. Si $E \subset \mathbb{R}^N$, $E = \{a_1, a_2, \ldots\}$, la σ-additivité de λ_N donne immédiatement que $\lambda_N(E) = 0$ si E est au plus dénombrable.

Il y a, en fait, beaucoup d'autres ensembles λ_N-négligeables dans $\mathbb{R}^N$. Toute courbe régulière simple dans $\mathbb{R}^N$ forme un ensemble λ_N-négligeable si $N \geqslant 2$; de même toute surface régulière dans $\mathbb{R}^N$, $N \geqslant 3$, est un ensemble λ_N-négligeable ; plus généralement, toute sous-variété différentielle de dimension k de classe C^1 dans $\mathbb{R}^N$, $N > k$, est un ensemble λ_N-négligeable (cf. sect. 5.5, aussi exercices § 5.4.5).

Si f, g sont deux fonctions $X \to \overline{\mathbb{R}}$, on dit que $f = g$ p.p.(μ) si $\{x : f(x) \neq g(x)\}$ est μ-négligeable. Pour illustrer ces notions donnons une proposition qui est souvent utile.

5.2.12 Proposition

Soit (X, Σ, μ) un espace mesuré.

(i) Si $f : X \to \overline{\mathbb{R}}$ est μ-intégrable alors f est fini p.p.(μ).

(ii) Si $f : X \to \overline{\mathbb{R}}$ est μ-intégrable et $g : X \to \overline{\mathbb{R}}$ est mesurable avec $g = f$ p.p.(μ) alors g est μ-intégrable.

DÉMONSTRATION

Commençons par établir une inégalité simple concernant les fonctions μ-intégrables f (appelé ***inégalité de Chebyshev***) : soit $A_\lambda = \{x : |f(x)| \geqslant \lambda\}$, $\lambda > 0$; alors

$$\mu(A_\lambda) \leqslant \frac{1}{\lambda} \int_{A_\lambda} |f| \, d\mu \leqslant \frac{1}{\lambda} \int |f| \, d\mu . \tag{5.13}$$

Cette inégalité est une conséquence de ce qui suit :

$$\int |f| \, d\mu \geqslant \int_{A_\lambda} |f| \, d\mu \geqslant \lambda \mu(A_\lambda)$$

car $|f| \cdot 1_{A_\lambda} \geqslant \lambda$.

(i) Posons $A_n = \{x \in X : |f(x)| \geqslant n\}$; d'après (5.13) $\mu(A_n) \leqslant M/n$ où $M = \int |f| \, d\mu$; si

$$A = \{x \in X : |f(x)| = \infty\} = \{x \in X : f(x) = \infty \text{ ou } -\infty\}$$

alors $A \subset A_n$, $n \geqslant 1$, de sorte que $\mu(A) = 0$ car

$$\mu(A) \leqslant \mu(A_n) \leqslant \frac{M}{n} \longrightarrow 0 \quad (n \to \infty) .$$

Le fait que $\mu(A) = 0$ équivaut à la phrase « f est fini p.p.(μ) ».

(ii) Il suffit de démontrer cette affirmation pour $f \geqslant 0$, $g \geqslant 0$, le cas général étant traîté par la considération de $f^{\pm}$, $g^{\pm}$.

Posons $A = \{x : f(x) = g(x) \in \mathbb{R}_+\}$; si $E = \{x : f(x) = g(x)\}$, $F = \{x : f(x) \in \mathbb{R}_+\}$, on sait de l'hypothèse que $\mu(E^c) = 0$ et, de (i), que $\mu(F^c) = 0$; puisque $A = E \cap F$, $A^c = E^c \cup F^c$ et on a

$$\mu(A^c) \leqslant \mu(E^c) + \mu(F^c) = 0 .$$

Si $h : X \to \overline{\mathbb{R}}$ est tel que $h(x) = 0$ si $x \in A$, et $h(x) = \infty$ si $x \in A^c$, on a que $\int h \, d\mu = 0$; puisque $g \leqslant f + h$ et $f \leqslant g + h$ on aura, successivement,

$$\int g \, d\mu \leqslant \int \cdot f \, d\mu , \quad \int f \, d\mu \leqslant \int g \, d\mu$$

d'où $\int f \, d\mu = \int g \, d\mu$. ♦

REMARQUES

(1) Dans cette proposition, on peut même considérer f, g comme μ-mesurables (au lieu de mesurables) quitte à remplacer Σ par Σ_μ et μ par son prolongement $\widetilde{\mu}$.

(2) La proposition 5.2.12 permet de parler de la μ-intégrabilité d'une fonction f dans A même lorsque f n'est définie que p.p.(μ) dans A ; il suffit de compléter la définition de f dans A en posant $f(x) = 0$ (ou une autre constante) aux points x de A où $f(x)$ n'est pas définie au départ. Selon la proposition 5.2.12, $\int_A f \, d\mu$ est indépendant de la définition de f sur une partie négligeable de A.

(3) Dans les énoncés des propositions 5.2.5, 5.2.8 et 5.2.9 on peut affaiblir les conditions au point de n'être valables que p.p.(μ). Indiquons deux de ces énoncés généralisés.

- Si $\{f_n(x)\}$ est une suite croissante p.p.(μ) et $f(x) = \lim_{n\to\infty} f_n(x)$ p.p.(μ) où $f_n \geqslant 0$ p.p.(μ), f_n μ-mesurables, $n \geqslant 1$, alors

$$\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu \, .$$

- Si $\{f_n\}$, $n = 1, 2, \ldots$, f_n sont μ-mesurables et μ-intégrables et $f_n(x) \to f(x)$, $n \to \infty$, $|f_n(x)| \leqslant M(x)$, p.p.(μ), M étant μ-intégrable, alors

$$\lim_{n\to\infty} \int |f_n - f| \, d\mu = 0 \, .$$

5.2.13 Intégrales de Lebesgue dans $\mathbb{R}^N$

Nous spécialisons toute la discussion des paragraphes précédents au cas de $X = \mathbb{R}^N$ et la mesure de Lebesgue λ_N est définie dans le paragraphe 5.2.2 ; elle est prolongée ensuite dans le paragraphe 5.2.11 par sa complétion Σ_{λ_N} que nous noterons ici par $\widetilde{\Sigma}$.

Les fonctions f qui sont λ_N-mesurables s'appellent ***mesurables*** ; si $f : \mathbb{R}^N \to \mathbb{R}$, ceci veut dire que l'on a $g \leqslant f \leqslant h$, g, h fonctions boréliennes telles que $g = h$ p.p.(λ_N) ; aussi on écrira simplement $g = h$ p.p., sans mentionner λ_N. Une fonction $f : \mathbb{R}^N \to \mathbb{R}$ est mesurable si et seulement si les ensembles $f^{-1}(B) \in \widetilde{\Sigma}$ pour B borélien dans $\mathbb{R}$; on peut définir $f : \mathbb{R}^N \to \overline{\mathbb{R}}$ comme étant mesurable si, en plus, $f^{-1}(\infty)$, $f^{-1}(-\infty)$ sont dans $\widetilde{\Sigma}$. Par « f ***intégrable*** » dans ce contexte on entendra une fonction $f : \mathbb{R}^N \to \mathbb{R}$ (ou $f : \mathbb{R}^N \to \overline{\mathbb{R}}$) qui est mesurable et λ_N-intégrable ; par $\mathcal{L}^1(\mathbb{R}^N)$ on indiquera l'ensemble de ces fonctions intégrables à valeurs dans $\mathbb{R}$.

Si $f \in \mathcal{L}^1(\mathbb{R}^N)$ on utilise l'écriture suivante :

$$\int f \, d\lambda_N = \int_{\mathbb{R}^N} f(x)\, dx = \int_{\mathbb{R}^N} f = \int f \, ;$$

si $A \subset \mathbb{R}^N$ est un sous-ensemble λ_N-mesurable, on dira que A est ***mesurable*** ce qui veut dire que l'on a $A_1 \subset A \subset A_2$, A_1, A_2 boréliens et $\lambda_N(A_2 \setminus A_1) = 0$; si $f : A \to \mathbb{R}$ est mesurable et λ_N-intégrable, on dira que f est ***intégrable dans*** A et on écrira

$$\int_A f \, d\lambda_N = \int_A f(x)\, dx = \int_A f \, .$$

Pour le cas de $A = [a, b]$ ou $[a, b[$ ou $]a, b]$ ou $]a, b[$ dans $\mathbb{R}$ on écrira

$$\int_A f \, d\lambda = \int_a^b f(x)\, dx = \int_a^b f$$

pourvu que f soit intégrable dans A. Cette écriture ne contredit pas l'écriture analogue dans le cas de l'intégrale de Riemann car on peut démontrer ce qui suit : *soit $f : [a, b] \to \mathbb{R}$ une fonction intégrable au sens de Riemann ; alors f est mesurable* (mais non nécessairement borélienne) *et intégrable ; aussi l'intégrale de Riemann de f sur $[a, b]$ coïncide avec celle de Lebesgue. Une fonction est intégrable au sens de Riemann si et seulement si elle est continue p.p.*

Par contre, on verra dans la suite beaucoup de fonctions qui sont intégrables (dans le sens de Lebesgue) sans être intégrables au sens de Riemann. Mais il faut comprendre que les intégrales dites ***impropres*** de Riemann peuvent exister sans que les fonctions correspondantes soient nécessairement intégrables au sens de Lebesgue. On reviendra à ce point dans la suite (cf. Ex. 21, § 5.2.15).

5.2.14 Proposition

Soit $f : K \to \mathbb{R}$ une fonction continue, K étant une partie compacte de $\mathbb{R}^N$. Pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que (en écrivant λ pour λ_N)

$$\left| \int_K f - \sum_{i=1}^{n} f(\xi_i)\lambda(A_i) \right| < \varepsilon$$

dès que les ensembles mesurables A_i satisfont les conditions suivantes avec $\xi_i \in A_i$, $i = 1, \ldots, n$: $A_i \subset K$, diamètre $(A_i) < \delta$, $i = 1, \ldots, n$; $\lambda(A_i \cap A_j) = 0$ si $i \neq j$; $\lambda\big(K \setminus \bigcup_{i=1}^n A_i\big) = 0$.

DÉMONSTRATION

Notons d'abord que $\lambda(K) < \infty$ car K, étant borné, est contenu dans un pavé borné dont la mesure est finie. Si $\lambda(K) = 0$, l'énoncé est banal car alors $\int_K f$ et $\lambda(A_i)$ sont tous nuls ; supposons donc que $\lambda(K) > 0$ et prenons un

ε' tel que $0 < \varepsilon' < \varepsilon/\lambda(K)$. Puisque f est uniformément continue dans K (à cause de la compacité de K), il existe un $\delta > 0$ tel que $|f(x) - f(y)| < \varepsilon'$ dès que $|x - y| < \delta$ avec x, y dans K. Posons

$$E = K \setminus \bigcup_{i=1}^{n} A_i\,, \quad B_1 = A_1\,,\ B_2 = A_2 \setminus A_1\,,\ B_3 = A_3 \setminus (A_1 \cup A_2)\,,\ \ldots$$

Alors les B_i sont deux à deux disjoints et

$$\lambda(E) = 0\,, \quad \lambda(A_i) = \lambda(B_i)\,,\ 1 \leqslant i \leqslant n\,, \quad \bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} B_i\,.$$

Si $g = \sum_{i=1}^{n} f(\xi_i)\, 1_{B_i}$, alors

$$\int_K g = \sum_{i=1}^{n} f(\xi_i)\lambda(B_i) = \sum_{i=1}^{n} f(\xi_i)\lambda(A_i)$$

et

$$\left|\int_K f - \int_K g\right| \leqslant \int_K |f - g| = \sum_{i=1}^{n} \int_{B_i} |f - g| \quad (\text{car } \lambda(E) = 0)$$

$$\leqslant \sum_{i=1}^{n} \int_{B_i} |f - f(\xi_i)| \leqslant \sum_{i=1}^{n} \varepsilon'\lambda(B_i) = \varepsilon'\lambda(K) < \varepsilon\,. \qquad \blacklozenge$$

REMARQUES. La proposition 5.2.14 montre, en particulier, que si $K = [a\,,\, b] \subset \mathbb{R}$, alors l'intégrale de Riemann d'une fonction continue f dans K est la même chose que l'intégrale de Lebesgue ; la proposition démontre même que pour le calcul de $\int_K f$ on peut choisir des « partitions » assez générales de K.

Dans les exercices qui suivent, on admettra comme connue toutes les règles usuelles de l'intégration des fonctions continues dans un intervalle compact.

5.2.15 Exercices

1 Soit (X, Σ, μ) un espace mesuré. Démontrer les affirmations suivantes directement en admettant que tous les ensembles concernés sont dans Σ :

(i) si $A \subset B$ alors $\mu(A) \leqslant \mu(B)$ et $\mu(A) + \mu(B \setminus A) = \mu(B)$;
(ii) $\mu(A) = \mu(A \setminus B) + \mu(A \cap B)$; $\mu(B) = \mu(B \setminus A) + \mu(A \cap B)$;
(iii) $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$;
(iv) $\mu(A \cup B) \leqslant \mu(A) + \mu(B)$;
(iv) $\mu(A_1 \cup A_2 \cup \cdots) \leqslant \sum_{i=1}^{\infty} \mu(A_i)$.

2 Soit $X = \{x_1, x_2, \ldots\}$ un ensemble au plus dénombrable. Si $p_i \geqslant 0$, $i \geqslant 1$, et $A \subset X$, posons

$$\mu(A) = \sum_{i=1}^{\infty} p_i \, 1_A(x_i) \, .$$

Vérifier que μ définit une mesure sur la tribu 2^X et que $\mu(X) < \infty$ si et seulement si $\sum_i p_i < \infty$. (On aura besoin du lemme suivant : si $a_{ij} \geqslant 0$, $i \geqslant 1$, $j \geqslant 1$, alors

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} \, .)$$

3 Soient X, μ comme dans l'exercice 2. Vérifier que : si $f : X \to \mathbb{R}$, alors f est mesurable ; si $f \geqslant 0$, $\int f \, d\mu = \sum_i f(x_i) p_i \leqslant \infty$; f est μ-intégrable si et seulement si $\sum_i |f(x_i)| p_i < \infty$ et alors $\int f \, d\mu = \sum_i f(x_i) p_i$. (Le cas $p_i = 1$, $i \geqslant 1$, donne la mesure de comptage sur X et alors $\int f \, d\mu$ donne la série $\sum_i f(x_i)$.)

4 Soit $X = \mathbb{N}^*$ muni de la mesure de comptage μ. Posons

$$A_n = \{1, \ldots, n\} \, , \quad f_n = (1/n) \, 1_{A_n} \, , \quad n = 1, 2, \ldots$$

Vérifier que $f_n \to 0$ uniformément dans X mais $\int f_n \, d\mu = 1$, $n \geqslant 1$; voir que $\{f_n\}$ ne vérifie pas la condition de domination dans le théorème de convergence dominée.

5 Soient $a_{nk} \geqslant 0$, $n, k \geqslant 1$ tels que $\sum_{k=1}^{\infty} a_{nk} = 1$ pour $n \geqslant 1$ et $\lim_{n \to \infty} a_{nk} = a_k$ pour $k \geqslant 1$. Démontrer que $\sum_{k=1}^{\infty} a_k \leqslant 1$ directement et aussi en utilisant le lemme de Fatou dans l'espace mesuré de l'exercice 4 ; donner un exemple ou $\sum_k a_k < 1$.

6 Dans l'exercice 5 voir que, sans la condition $a_{nk} \geqslant 0$, on peut avoir des exemples où $\sum_k a_{nk} = 1$, $\lim_{n \to 0} a_{nk} = a_k$ pour $k \geqslant 1$ mais $\sum_k a_k > 1$ (essayer $a_{n1} = 2$, $a_{n(n+1)} = -1$, $a_{nj} = 0$ si $j \neq 1, n+1$). Quelle relation y a-t-il avec le lemme de Fatou ?

7 Soit A un sous-ensemble dénombrable dans $\mathbb{R}^N$; posons $f = 1_A$. Voir que $\int f(x) \, dx = 0$. En particulier, si $A = \mathbb{Q} \subset \mathbb{R}$, $f = 1_A$ est une fonction discontinue. (Ceci donne un exemple d'une fonction intégrable au sens de Lebesgue sans être intégrable au sens de Riemann sur un intervalle compact quelconque.)

8 Soit $f(x) = 0$ si x est rationnel et $f(x) = x$ si x est irrationnel, $0 \leqslant x \leqslant 1$. Vérifier que $\int_0^1 f(x) \, dx = 1/2$.

9 Soit $f(x) = x^\alpha$, $x > 0$ ($\alpha \in \mathbb{R}$). Démontrer que f est intégrable dans $[0, 1]$ si et seulement si $\alpha > -1$. Démontrer que f est intégrable dans $[1, \infty[$ si et seulement si $\alpha < -1$. (Calculer $\int_\varepsilon^1 f$, $\int_1^T f$ et étudier $\varepsilon \to 0$, $T \to \infty$.)

10 Soit $f(x) = x^\alpha |\ln x|^\beta$, $x > 0$ (α, β réels). Démontrer que :

(i) f est intégrable sur $[0, a]$ ($0 < a < 1$) si et seulement si $\alpha > -1$ (et β quelconque) ou $\alpha = -1$ et $\beta < -1$;

(ii) f est intégrable sur $[a, b]$ ($0 < a \leqslant 1$, $1 < b < \infty$) si et seulement si $\beta > -1$ (et α quelconque) ;

(iii) f est intégrable sur $[a, \infty[$ ($a > 1$) si et seulement si $\alpha < -1$ (et β quelconque) ou $\alpha = -1$ et $\beta < -1$.

11 Soit $f : [0, \infty[\to \mathbb{R}_+$ défini comme suit : on prend $0 < a_n < 1$, $n = 1, 2, \ldots$ tels que $\sum_{n=1}^\infty n a_n < \infty$; on pose $f(n) = f(n + a_n) = 0$, $f(n + a_n/2) = 2n$ et on définit $f \geqslant 0$ dans $[n, n + a_n]$ de telle manière que le graphe de $f|_{[n, n+a_n]}$ est un triangle isocèle de base de longueur a_n et hauteur $2n$; posons $f(x) = 0$ si x n'est pas dans un des intervalles $[n, n + a_n]$, $n \geqslant 1$. Faire un dessin de f ; vérifier que $\int_0^\infty f(x)\,dx = \sum_n n a_n < \infty$, f est continue mais $\lim_{x\to\infty} f(x)$ n'existe pas.

12 Soit (X, Σ, μ) un espace mesuré. Si $A_n \in \Sigma$, $n = 1, 2, \ldots$, et $A_1 \supset A_2 \supset \ldots$, $A = \bigcap_{n=1}^\infty A_n$, alors $\mu(A) = \lim_{n\to\infty} \mu(A_n)$ si $\mu(A_p) < \infty$ pour un $p \geqslant 1$ (cf. prop. 5.2.6(vi)).

13 Soit (X, Σ, μ) un espace mesuré et $f \in \mathcal{L}^1(\mu)$. Démontrer que

$$\lim_{n\to\infty} \int_{A_n} f\,d\mu = 0$$

si $A_n \in \Sigma$, $n = 1, 2, \ldots$, $A_1 \supset A_2 \supset \ldots$, $\bigcap_{n=1}^\infty A_n = \emptyset$.

14 Soit $f : [0, \infty[\to \mathbb{R}_+$ une fonction continue. Vérifier que f est intégrable si et seulement si pour tout $\varepsilon > 0$ il existe $T_\varepsilon > 0$ tel que $\int_t^{t'} f < \varepsilon$ dès que $T_\varepsilon \leqslant t < t' < \infty$. En déduire que si $\lim_{x\to\infty} f(x) = a$ et f est intégrable sur $[0, \infty[$, alors $a = 0$.

15 Soit $f_n : [0, 1] \to \mathbb{R}_+$ défini par $f_n = a_n 1_{]0, 1/n[}$, $a_n > 0$, $n = 1, 2, \ldots$ Vérifier que $\lim_{n\to\infty} f_n(x) = 0$ pour tout $x \in [0, 1]$ et $\int_0^1 f_n = a_n/n$; en variant le choix des a_n on obtient $f_n \to 0$ ($n \to \infty$), mais où la limite de $\int_0^1 f_n$ est soit différente de 0, soit inexistante.

16 On sait que

$$(1+x)^{-1} = 1 - x + x^2 - \cdots, \quad 0 \leqslant x < 1$$
$$(1+x^2)^{-1} = 1 - x^2 + x^4 - \cdots, \quad 0 \leqslant x < 1.$$

En intégrant terme à terme, obtenir les formules suivantes

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \cdots ; \quad \frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots$$

Justifier ces calculs en utilisant le théorème de convergence dominée (par exemple $\left|1 - x + \cdots + (-1)^n x^n\right| \leqslant 2/(1+x)$, $0 \leqslant x < 1$, cf. la formule de la somme des séries géométriques).

17 De (cf. Ex. 16) :

$$\frac{1}{2} \ln\left(\frac{1+x}{1-x}\right) = x + \frac{x^3}{3} + \cdots, \quad |x| < 1$$

obtenir

$$\frac{1}{2}\int_0^1 \ln\left(\frac{1+x}{1-x}\right) \mathrm{d}x = \sum_{n=1}^{\infty} \frac{1}{2n(2n-1)}$$
$$\frac{1}{2}\int_0^1 \frac{1}{x} \ln\left(\frac{1+x}{1-x}\right) \mathrm{d}x = 2\left(1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots\right) ;$$

la justification à donner est celle de l'intégration des séries à termes $\geqslant 0$. (On sait que $1 + 1/3^2 + 1/5^2 + \cdots = \pi^2/8$.)

18 (Fonction gamma) Soit $f_t(x) = \mathrm{e}^{-x} x^{t-1}$, $x > 0$ ($t > 0$). Vérifier que f_t est intégrable sur $[0, \infty[$ (considérer séparément les intervalles $[0, 1]$ et $[1, \infty[$). Posons

$$\Gamma(t) = \int_0^{\infty} f_t(x)\, \mathrm{d}x .$$

Vérifier que : $\Gamma(1) = 1$; $\Gamma(n) = (n-1)!$ si $n \in \mathbb{N}^*$; $\Gamma(t+1) = t\Gamma(t)$, $t > 0$ (en faisant une intégration par parties).

19 (Fonction beta) Soit $f_{s,t}(x) = x^{s-1}(1-x)^{t-1}$, $0 < x < 1$ ($s > 0, t > 0$). Vérifier que $f_{s,t}$ est intégrable dans $[0, 1]$. Posons

$$B(s,t) = \int_0^1 x^{s-1}(1-x)^{t-1}\, \mathrm{d}x .$$

Vérifier que $B(s,t) = B(t,s)$; aussi

$$\int_0^{\pi/2} (\sin\theta)^{\alpha} (\cos\theta)^{\beta}\, \mathrm{d}\theta = \frac{1}{2} B\left(\frac{\alpha+1}{2}, \frac{\beta+1}{2}\right)$$

pour $\alpha > -1$, $\beta > -1$. (On verra que $B(s,t) = \Gamma(s)\Gamma(t)/\Gamma(s+t)$).

20 Vérifier que $\int_0^1 x^{-x}\,dx = \sum_{n=1}^{\infty} n^{-n}$ en écrivant

$$x^{-x} = \exp(-x\ln x) = \sum_{m=0}^{\infty} \frac{(-x\ln x)^m}{m!}$$

et en intégrant terme à terme (vous pouvez utiliser la fonction gamma de l'exercice 18).

21 Vérifier que

$$\lim_{T\to\infty} \int_{\pi/2}^{T} \frac{\sin x}{x}\,dx = -\int_{\pi/2}^{\infty} \frac{\cos x}{x^2}\,dx$$

en faisant une intégration par parties. En déduire que $\lim_{T\to\infty} \int_{\pi/2}^{T} \sin x/x\,dx$ existe, ce que l'on décrit en disant que $\int_0^{\infty} \sin x/x\,dx$ existe en tant qu'une intégrale impropre. Démontrer que

$$\int_0^{\infty} \left|\frac{\sin x}{x}\right| dx = \infty$$

en vérifiant les affirmations suivantes :

$$\int_0^{\infty} \left|\frac{\sin x}{x}\right| dx = \sum_{n=0}^{\infty} \int_0^{\pi} \frac{\sin y}{n\pi + y}\,dy = \int_0^{\pi} \sum_{n=0}^{\infty} \frac{\sin y}{n\pi + y}\,dy = \infty$$

car

$$\sum_{n=0}^{\infty} \frac{\sin y}{n\pi + y} = \infty, \quad 0 < y < \pi.$$

(Noter que l'on a ici des fonctions positives dans $[0, \pi]$.)

22 Soit (X, Σ, μ) un espace mesuré. Si $A \in \Sigma$ et $f : A \to \overline{\mathbb{R}}$ est μ-intégrable avec $\int_A |f|\,d\mu = 0$, alors $f = 0$ p.p.(μ) dans A ; si $X = \mathbb{R}^N$, $\mu = \lambda_N$ et f est de plus continue dans A, supposé ouvert dans $\mathbb{R}^N$, alors $f \equiv 0$ dans A.
(Si $A_n = \{x \in A : |f(x)| > 1/n\}$, on voit de l'inégalité de Chebyshev (5.13) que $\mu(A_n) = 0$ d'où $\mu\big(\bigcup_{n=1}^{\infty} A_n\big) = 0$ ce qui équivaut à avoir $f = 0$ p.p.(μ) dans A. Si f est continue dans A et $f(a) \neq 0$ pour un $a \in A$, alors il existe un pavé ouvert $P \subset A$, P contenant a, et un $\varepsilon > 0$ tel que $|f(x)| > \varepsilon$, $x \in P$; alors $\int_A |f| \geqslant \varepsilon\lambda_N(P) > 0$ donne une contradiction.)

23 Soit $f_n(x) = n^2 x(1-x)^n$, $n = 1, 2, \ldots$; posons $\alpha_n = \int_0^1 f_n(x)\,dx$. Vérifier que $\alpha_n = n^2/(n+1)(n+2) \to 1$ lorsque $n \to \infty$ mais que $\lim_{n\to\infty} f_n(x) = 0$, $0 \leqslant x \leqslant 1$. Voir explicitement que la condition de domination du théorème de convergence dominée est en défaut.

24 $\int_0^\infty x/(e^x - 1)\,dx = \pi^2/6$ (en admettant que $\sum_{n=1}^\infty n^{-2} = \pi^2/6$). (Ramener l'intégrale à $-\int_0^1 (1/y)\ln(1-y)\,dy$, développer $-\ln(1-y)$ en série.)

25 Soit (X, Σ, μ) un espace mesuré. Si $p : X \to [0\ \infty]$ est une fonction μ-intégrable, posons

$$\nu_p(A) = \int_A p\,d\mu, \quad A \in \Sigma.$$

Vérifier que ν_p est une mesure positive finie.

26 Soit (X, Σ, μ) un espace mesuré σ-fini ; si $p : X \to [0\ \infty[$ est mesurable, définissons ν_p comme dans l'exercice 25. Démontrer que ν_p est une mesure σ-finie. (Considérer $E_j = \{x : p(x) \leqslant j\}$, $A_{nj} = A_n \cap E_j$ où les A_n sont comme dans la définition de σ-finitude de μ.)

27 Soit (X, Σ, μ) un espace mesuré ; si $p : X \to [0, \infty]$ est mesurable, définissons ν_p comme dans l'exercice 25. Vérifier que ν_p est une mesure positive telle que $\nu_p(A) = 0$ si $\mu(A) = 0$. Si $f : X \to [0, \infty]$ est mesurable, démontrer que

$$\int f\,d\nu_p = \int fp\,d\mu.$$

Une fonction mesurable $f : X \to \overline{\mathbb{R}}$ est ν_p-intégrable si et seulement si $f \cdot p$ est μ-intégrable. (Vérifier ces affirmations pour $f = 1_A$; cf. § 5.2.10.)

REMARQUE. La fonction p dans les exercices 25, 26 et 27 s'appelle la ***densité*** de la mesure positive ν_p par rapport à la mesure positive μ. Un théorème important (appelé le ***théorème de Lebesgue-Radon-Nikodym***) affirme que si μ est une mesure σ-finie sur Σ alors toute mesure positive ν sur Σ, telle que $\nu(A) = 0$ dès que $\mu(A) = 0$, est de la forme ν_p pour une densité appropriée $p : X \to [0, \infty]$. La condition « $\nu(A) = 0$ si $\mu(A) = 0$ » est généralement décrite par la phrase « ν est ***absolument continue*** par rapport à μ ».

5.3 INTÉGRALES QUI DÉPENDENT D'UN PARAMÈTRE

Soit (X, Σ, μ) un espace mesuré. Un problème qui se présente souvent en analyse est l'étude d'une fonction $g : T \to \mathbb{R}$, $T \subset \mathbb{R}^M$, définie comme suit :

$$g(t) = \int_X f(t, x)\mu(dx), \quad t \in T,$$

où l'on supose que, pour tout $t \in T$, la fonction $x \mapsto f(t, x)$ est μ-intégrable. Ici, on va répondre à la question de savoir sous quelles conditions sur f, la fonction g sera continue ou dérivable.

Pour toutes les propositions (et corollaires) de cette section, la notation établie ci-dessus restera en vigueur.

5.3.1 Proposition

Soit $\tau \in \overline{T}$; supposons qu'il existe un voisinage ouvert V de τ tel que

$$\left|f(t,x)\right| \leqslant M(x)\,, \quad t \in V \cap T\,,\ x \in X\,,$$

où $M : X \to \overline{\mathbb{R}}_+$ est μ-intégrable.

Si $\lim_{t\to\tau} f(t,x) = \varphi(x)$ p.p.(μ) alors φ est μ-intégrable et

$$\lim_{t\to\tau} g(t) = \int_X \varphi(x)\mu(\mathrm{d}x)\,.$$

En particulier, si $t \mapsto f(t,x)$ est continue au point τ pour x p.p.(μ) (c'est-à-dire, $\varphi(x) = f(\tau,x)$ p.p.(μ)), alors g est continue au point τ.

Si T est non borné et si on suppose que

$$\left|f(t,x)\right| \leqslant M(x)\,, \quad \left\{t \in T : |t| > R\right\},\ x \in X\,,$$

pour un $R > 0$ et $M : X \to \overline{\mathbb{R}}$ μ-intégrable et

$$\lim_{|t|\to\infty} f(t,x) = \varphi(x) \quad \text{p.p.}(\mu)\,,$$

alors φ est μ-intégrable et

$$\lim_{|t|\to\infty} g(t) = \int_X \varphi(x)\mu(\mathrm{d}x)\,.$$

Si T est un intervalle non borné dans $\mathbb{R}$, on a des énoncés analogues pour $\lim_{t\to\infty} g(t)$, $\lim_{t\to-\infty} g(t)$.

DÉMONSTRATION

Si $\{t_n\}$, $n = 1, 2, \ldots$, est une suite dans T qui converge vers τ, on aura $t_n \in V \cap T$ pour $n \geqslant n_0$; le théorème de convergence dominée donnera alors que $\lim_{n\to\infty} g(t_n)$ existe et est égale à

$$\int_X \varphi(x)\mu(\mathrm{d}x)\ ;$$

ceci étant vrai pour toute suite $\{t_n\}$ convergeant par τ, on a le résultat donné pour $\lim_{t\to\tau} g(t)$. Les démonstrations des autres affirmations de la proposition sont tout à fait similaires. ♦

5.3.2 Corollaire

Supposons que $X \subset \mathbb{R}^N$, $\mu = \lambda_N$, et que X, T sont compacts avec $f : T \times X \to \mathbb{R}$ continue. Alors g est continue dans T.

DÉMONSTRATION

Il suffit de remarquer que

$$\left|f(t,x)\right| \leqslant M\,, \quad (t,x) \in T \times X\,,$$

pour un $M < \infty$ approprié, à cause de la continuité de f sur l'ensemble compact $T \times X \subset \mathbb{R}^{M+N}$ et que $\lambda_N(X) < \infty$ car X est compact dans $\mathbb{R}^N$. ♦

5.3.3 Proposition (*dérivation sous le signe d'intégration*)

Supposons que T est une partie ouverte de $\mathbb{R}^M$ et que, pour $x \in E^c$ avec $\mu(E) = 0$, $\mathrm{D}_i f(t,x)$ existe pour $t \in T$, $t = (t_1, \ldots, t_M)$, où D_i signifie $\partial/\partial t_i$, i étant un indice fixe parmi $\{1, \ldots, M\}$. Supposons qu'il existe $M : X \to \overline{\mathbb{R}}$ μ-intégrable tel que

$$\left|\mathrm{D}_i f(t,x)\right| \leqslant M(x)\,, \quad t \in T\,,\ x \in E^c\,.$$

Alors $\mathrm{D}_i g(t)$ existe pour $t \in T$ et

$$\mathrm{D}_i g(t) = \int_X \mathrm{D}_i f(t,x)\mu(\mathrm{d}x)\,, \quad t \in T\,, \tag{5.14}$$

Si T est un intervalle quelconque dans $\mathbb{R}$ et si $\mathrm{D}_t f(t,x)$ représente la dérivée partielle de f par rapport à t (dérivée unilatérale si t est une extrémité finie de T) alors (5.14) reste valable avec D_t au lieu de D_i.

DÉMONSTRATION

Soit B une boule ouverte contenant t telle que $\overline{B} \subset T$; posons e_i pour le i-ième vecteur de la base canonique de $\mathbb{R}^M$; alors pour t et $t + he_i$ dans B (h réel, $h \neq 0$) on a

$$\frac{1}{h}\left\{g(t+he_i) - g(t)\right\} = \int_X \frac{1}{h}\left\{f(t+he_i\,,\,x) - f(t,x)\right\}\mu(\mathrm{d}x)\,.$$

Puisque

$$\left|\frac{1}{h}\left\{f(t+he_i\,,\,x) - f(t,x)\right\}\right| = \left|\mathrm{D}_i f(s,x)\right| \leqslant M(x)\,, \quad x \in E^c\,,$$

où s est un point de B sur le segment $[\,t\,,\,t+he_i\,]$, et

$$\lim_{h\to 0}\frac{1}{h}\left\{f(t+he_i\,,\,x) - f(t,x)\right\} = \mathrm{D}_i f(t,x)\,, \quad x \in E^c\,;$$

on obtient (5.14) en se servant de la proposition 5.3.1 et du fait que $\mu(E) = 0$. Le cas d'un intervalle $T \subset \mathbb{R}$ est complètement analogue. ♦

4 De la formule élémentaire

$$\int_0^\infty \frac{\mathrm{d}x}{x^2+t^2} = \frac{\pi}{2t}, \quad t>0,$$

obtenir

$$\int_0^\infty \frac{\mathrm{d}x}{(x^2+t^2)^n} = \frac{\pi}{2}\cdot\frac{1\cdot 3\cdots(2n-3)}{2\cdot 4\cdots(2n-2)}\cdot\frac{1}{t^{2n-1}}, \quad t>0,$$

$n = 2, 3, \ldots$, par dérivations successives par rapport à t dans la formule élémentaire.

5 Comme dans l'exercice 4, des formules élémentaires

$$\int_0^\infty \mathrm{e}^{-tx}\sin x\,\mathrm{d}x = \frac{1}{1+t^2}, \quad \int_0^\infty \mathrm{e}^{-tx}\cos x\,\mathrm{d}x = \frac{t}{t^2+1} \quad (t>0)$$

déduire (avec des justifications) que

$$\int_0^\infty \mathrm{e}^{-tx}x\sin x\,\mathrm{d}x = \frac{2t}{(1+t^2)^2}, \quad \int_0^\infty \mathrm{e}^{-tx}x\cos x\,\mathrm{d}x = \frac{t^2-1}{(t^2+1)^2}$$

pour $t>0$.

6 De même, établir successivement les formules suivantes ($a>0$, $b>0$) :

(i) $\displaystyle\int_0^{\pi/2} \frac{\mathrm{d}x}{a^2\sin^2 x + b^2\cos^2 x} = \frac{\pi}{2ab}$ (intégrale élémentaire) ;

(ii) $\displaystyle\int_0^{\pi/2} \frac{\sin^2 x\,\mathrm{d}x}{(a^2\sin^2 x + b^2\cos^2 x)^2} = \frac{\pi}{4a^3b}$;

(iii) $\displaystyle\int_0^{\pi/2} \frac{\cos^2 x\,\mathrm{d}x}{(a^2\sin^2 x + b^2\cos^2 x)^2} = \frac{\pi}{4ab^3}$;

(iv) $A = \displaystyle\int_0^{\pi/2} \frac{\sin^2 x\,\mathrm{d}x}{a^2\sin^2 x + b^2\cos^2 x} = \frac{\pi}{2a(a+b)}$;

(v) $B = \displaystyle\int_0^{\pi/2} \frac{\cos^2 x\,\mathrm{d}x}{a^2\sin^2 x + b^2\cos^2 x} = \frac{\pi}{2b(a+b)}$

(en vérifiant que $A+B = \pi/(2ab)$, $a^2A + b^2B = \pi/2$) ;

(vii) $\displaystyle\int_0^{\pi/2} \ln(a^2\sin^2 x + b^2\cos^2 x)\,\mathrm{d}x = \pi\ln\Big(\frac{a+b}{2}\Big)$;

7 Soit

$$g(t) = \int_0^\infty \ln\Big(1 + \frac{t^2}{x^2}\Big)\,\mathrm{d}x\,.$$

Démontrer que $g(t)$ existe pour $t \in \mathbb{R}$ et que $g(0) = 0$. Pour $\alpha > 0$, poser $g_\alpha(t) = \int_0^\alpha \ln(1 + t^2/x^2)\,\mathrm{d}x$; vérifier que $g'_\alpha(t) = 2\operatorname{arc\,tg}(\alpha/t)$ pour $t > 0$ d'où

$$g_\alpha(t) = 2\int_0^t \operatorname{arc\,tg}\frac{\alpha}{u}\,\mathrm{d}u\,, \quad t > 0\,;$$

en conclure que

$$g(t) = \lim_{\alpha\to\infty} g_\alpha(t) = 2t\cdot\frac{\pi}{2} = \pi t\,, \quad t > 0$$

de sorte que

$$\int_0^\infty \ln\left(1 + \frac{t^2}{x^2}\right)\mathrm{d}x = \pi\cdot|t|\,, \quad t \in \mathbb{R}.$$

8 Soit $g(t) = \int_{-1}^1 f(x,t)\,\mathrm{d}x$, $t \in \mathbb{R}$, où

$$f(x,t) = \frac{\sin t}{x^2 - 2x\cos t + 1}\,, \quad -1 < x < 1\,.$$

Un calcul élémentaire montre que

$$\int_{-1}^1 \frac{\mathrm{d}x}{x^2 - 2x\alpha + 1} = \frac{\pi/2}{\sqrt{1-\alpha^2}}\,, \quad |\alpha| < 1$$

(écrire $x^2 - 2x\alpha + 1 = (x-\alpha)^2 + (1-\alpha^2)$; noter que, $\operatorname{arc\,tg} p + \operatorname{arc\,tg} 1/p = \pi/2$ si $p > 0$). Conclure que

$$g(t) = \begin{cases} \dfrac{\pi}{2} & \text{si } 2n\pi < t < (2n+1)\pi \\ -\dfrac{\pi}{2} & \text{si } (2n-1)\pi < t < 2n\pi \\ 0 & \text{si } t = n\pi \end{cases} \qquad (n = 0, \pm1, \pm2, \ldots)\,.$$

(Noter que f est de classe C^∞ dans $\mathbb{R} \times]-1, 1[$, mais que g est discontinue, montrant que les propositions 5.3.1 et 5.3.3 ne sont pas applicables.)

9 Soit $g(t) = \int_0^\infty (\sin tx/x)\,\mathrm{d}x$; alors

$$g(t) = \begin{cases} \dfrac{\pi}{2} & \text{si } t > 0 \\ 0 & \text{si } t = 0 \\ -\dfrac{\pi}{2} & \text{si } t < 0. \end{cases}$$

(Ici aussi g est discontinue ; la proposition 5.3.1 est inutilisable car il s'agit ici d'une intégrale impropre.)

5.4 THÉORÈME DE FUBINI

Dans cette section nous donnons un théorème général d'une application très commode qui permet de calculer une intégrale multiple comme une succession d'intégrales simples. De nouveau, pour des raisons de commodité de formulation, nous présentons le théorème d'abord dans un cadre général ; nous le spécialisons ensuite facilement au cas de l'intégration dans $\mathbb{R}^N$.

5.4.1 Définitions préliminaires

Soient $(X, \mathcal{F})$, $(Y, \mathcal{G})$ deux espaces boréliens ; un ensemble de la forme $A \times B \subset X \times Y$, où $A \in \mathcal{F}$, $B \in \mathcal{G}$, s'appellera un ***rectangle mesurable*** dans $X \times Y$; la tribu engendrée (cf. § 5.2.1) par les rectangles mesurables dans $X \times Y$ sera notée Σ ; cette tribu s'appelle la ***tribu produit*** de $\mathcal{F}$ et $\mathcal{G}$, et souvent on la note comme

$$\Sigma = \mathcal{F} \otimes \mathcal{G} .$$

Rappelons que Σ est la plus petite tribu dans $X \times Y$ contenant les rectangles mesurables $A \times B$, $A \in \mathcal{F}$, $B \in \mathcal{G}$.

Si $X = \mathbb{R}^m$ est muni de sa tribu borélienne Σ_m, et $Y = \mathbb{R}^n$ est muni de sa tribu borélienne Σ_n, alors on peut démontrer que *la tribu produit de* Σ_m *et* Σ_n *dans* $X \times Y = \mathbb{R}^{m+n}$ *est la tribu borélienne* Σ_{m+n} *dans* $\mathbb{R}^{m+n}$ autrement dit

$$\Sigma_{m+n} = \Sigma_m \otimes \Sigma_n .$$

Revenons au cas général du départ. Si $E \subset X \times Y$ et $x \in X$, $y \in Y$, posons

$$E_x = \{t \in Y : (x,t) \in E\}, \quad E^y = \{t \in X : (t,y) \in E\} ;$$

l'ensemble $E_x \subset Y$ s'appelle la ***x-section*** de E et l'ensemble $E^y \subset X$ s'appelle la ***y-section*** de E.

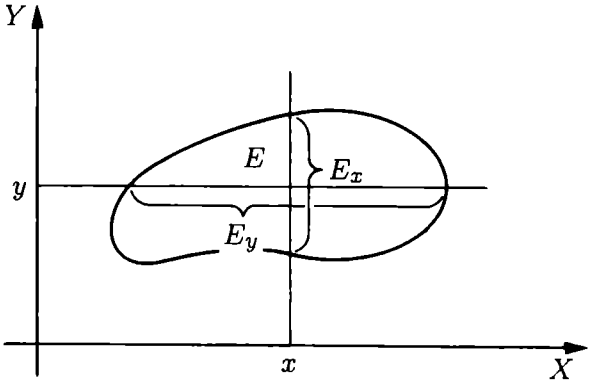


Fig. 5.2

Si $X = Y = \mathbb{R}$ la terminologie « x-section » « y-section » correspond aux notations géométriques courantes.

Une proposition principale du sujet est la suivante.

5.4.2 Proposition (*existence de la mesure produit*)

Soient $(X, \mathcal{F}, \mu)$, $(Y, \mathcal{G}, \nu)$ deux espaces mesurés σ-finis ; soit $\Sigma = \mathcal{F} \otimes \mathcal{G}$ la tribu produit dans $X \times Y$. Si $E \in \Sigma$ alors, pour tout $x \in X$, $E_x \in \mathcal{G}$ et pour tout $y \in Y$, $E^y \in \mathcal{F}$. Les fonctions positives $x \mapsto \nu(E_x)$, $y \mapsto \mu(E^y)$ sont mesurables et on a l'égalité suivante pour tout $E \in \Sigma$:

$$\rho(E) = \int_X \nu(E_x)\mu(\mathrm{d}x) = \int_Y \mu(E^y)\nu(\mathrm{d}y)\,. \tag{5.16}$$

De plus, $\rho : \Sigma \to [0, \infty]$ est une mesure σ-finie (appelée la ***mesure produit*** *de μ et ν ; on écrit aussi, $\rho = \mu \otimes \nu$). La mesure ρ est uniquement déterminée par la propriété que $\rho(A \times B) = \mu(A)\nu(B)$, $A \in \mathcal{F}$, $B \in \mathcal{G}$.*

DÉMONSTRATION (esquisse)

Si E est un rectangle mesurable $A \times B$ on vérifie immédiatement que $E_x = B$ ou $\emptyset$ selon que $x \in A$ ou non, et $E^y = A$ ou $\emptyset$ selon que $y \in B$ ou non ; dans ce cas (5.16) devient simplement

$$\rho(E) = \rho(A \times B) = \nu(B)\mu(A) = \mu(A)\nu(B)\,.$$

Le reste de la démonstration est un peu technique ; il consiste à vérifier que (5.16) reste vrai d'abord pour les réunions finies des rectangles mesurables puis pour la tribu engendrée Σ. C'est ici que la σ-finitude de μ et ν nous est utile. ♦

Si $\mu = \lambda_m$, $X = \mathbb{R}^m$, $\mathcal{F} = \Sigma_m$, $\nu = \lambda_n$, $Y = \mathbb{R}^n$, $\mathcal{G} = \Sigma_n$, Σ_m, Σ_n étant les tribus boréliennes correspondantes, on peut démontrer que λ_{m+n} *sur* $\Sigma_{m+n} = \Sigma_m \otimes \Sigma_n$ *dans* $\mathbb{R}^{m+n}$ *est la mesure produit de* λ_m *et* λ_n ; en particulier, si $E \in \Sigma_{m+n}$,

$$\lambda_{m+n}(E) = \int_{\mathbb{R}^m} \lambda_n(E_x)\,\mathrm{d}x = \int_{\mathbb{R}} \lambda_m(E^y)\,\mathrm{d}y\,. \tag{5.17}$$

La formule (5.17) est à la base des calculs des volumes des ensembles dans $\mathbb{R}^3$ (ou dans $\mathbb{R}^N$) en calculant les aires ou les longueurs des sections (ou leurs analogues dans les dimensions supérieures) et en les intégrant dans la suite. Nous nous référons à (5.17) en l'appelant le ***principe de Cavalieri***.

Le théorème suivant généralise la proposition 5.4.2 ; en fait, il en est une conséquence immédiate grâce à la théorie de l'intégration générale développée dans la section 5.2. Néanmoins, c'est le théorème central dans les applications des intégrales multiples.

Si $f : X \times Y \to \mathbb{R}$, $x \in X$, $y \in Y$, nous notons par f_x la fonction $t \mapsto f(x,t)$, $t \in Y$ et par f^y la fonction $t \mapsto f(t,y)$, $t \in X$; f_x, f^y s'appellent les ***fonctions partielles*** définies par f ; on a

$$f_x : Y \to \mathbb{R}, \quad f^y : X \to \mathbb{R}, \quad f_x(y) = f(x,y) = f^y(x).$$

5.4.3 Proposition (théorème de Fubini)

Soient $(X, \mathcal{F}, \mu)$, $(Y, \mathcal{G}, \nu)$ deux espaces mesurés σ-finis et ρ la mesure produit $\mu \otimes \nu$ définie sur la tribu produit $\Sigma = \mathcal{F} \otimes \mathcal{G}$ dans $X \times Y$.

Si $f : X \times Y \to \mathbb{R}$ est Σ-mesurable alors, pour tout $x \in X$, f_x est $\mathcal{G}$-mesurable et pour tout $y \in Y$, f^y est $\mathcal{F}$-mesurable.

Soit $f : X \times Y \to \overline{\mathbb{R}}_+$ une fonction numérique positive qui est Σ-mesurable ; alors

$$\int_{X\times Y} f \, d\rho = \int_X \left\{ \int_Y f(x,y)\, \nu(dy) \right\} \mu(dx) = \int_Y \left\{ \int_X f(x,y)\, \mu(dx) \right\} \nu(dy) \tag{5.18}$$

où les fonctions

$$x \longmapsto \int_Y f(x,y)\, \nu(dy), \quad y \longmapsto \int_X f(x,y)\, \mu(dx) \tag{5.19}$$

sont, respectivement, $\mathcal{F}$ et $\mathcal{G}$ mesurables.

Soit $f : X \times Y \to \mathbb{R}$ une fonction Σ-mesurable et ρ-intégrable ; alors (5.18) reste valable. Les deux fonctions de (5.19) existent p.p. dans le sens que pour x p.p.(μ), f_x est ν-intégrable et pour y p.p.(ν), f^y est μ-intégrable.

DÉMONSTRATION (esquisse)

Les énoncés sont évidemment valables pour $f = 1_E$, $E \in \Sigma$, grâce à la proposition 5.4.2 ; ils sont donc valables pour $f \geqslant 0$, simple et mesurable. Un passage à la limite suivant le schéma du paragraphe 5.2.10 donne les résultats donnés jusqu'à (5.18) et (5.19) pour $f \geqslant 0$, mesurable. La dernière partie se fait en décomposant f comme $f^+ - f^-$. ♦

REMARQUES

(1) Toute la discussion précédente se généralise sans difficulté au cas du produit de n espaces mesurés σ-finis $(X_i, \mathcal{F}_i, \mu_i)$, $i = 1, \ldots, n$, mais nous n'aurons pas besoin de cette généralisation.

(2) Même si les mesures μ, ν sont complètes, la mesure produit ρ n'est pas nécessairement complète. Si l'on considère $f : X \times Y \to \mathbb{R}$ qui est Σ_ρ-mesurable (au lieu de Σ-mesurable) dans la proposition 5.4.3, alors (5.18) reste valable avec la seule modification technique que f_x et f^y sont mesurables seulement p.p.(μ) et p.p.(ν) respectivement. Cette modification ne change pas grand chose dans les applications courantes.

5.4.4 Cas de $\mathbb{R}^N$

Nous spécialisons la proposition 5.4.3 au cas où

$$X = \mathbb{R}^m\,,\ \mathcal{F} = \Sigma_m\,,\ \mu = \lambda_m\,;\quad Y = \mathbb{R}^n\,,\ \mathcal{G} = \Sigma_n\,,\ \nu = \lambda_n\,.$$

Posons $m + n = N$ et considérons $\mathbb{R}^N$ comme $X \times Y$; on a déjà vu ci-dessus que

$$\Sigma_N = \Sigma_m \otimes \Sigma_n\,,\quad \lambda_N = \lambda_m \otimes \lambda_n\,.$$

Ecrivons un point de $\mathbb{R}^N$ comme $z = (x, y)$, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$; alors la formule (5.18) devient ce qui suit :

$$\int_{\mathbb{R}^N} f(z)\,\mathrm{d}z = \int_{\mathbb{R}^m} \left\{ \int_{\mathbb{R}^n} f(x,y)\,\mathrm{d}y \right\} \mathrm{d}x = \int_{\mathbb{R}^n} \left\{ \int_{\mathbb{R}^m} f(x,y)\,\mathrm{d}x \right\} \mathrm{d}y \quad (5.20)$$

sous condition que $f : \mathbb{R}^N \to \mathbb{R}$ est borélienne et $\geqslant 0$ ou intégrable. La condition que f est borélienne ne sera jamais nécessaire à vérifier dans ce livre ; donc $f \geqslant 0$ (même p.p.) serait une condition suffisante pour garantir (5.20). Si f est de signe variable, il est essentiel de vérifier que f est intégrable ; ceci est souvent facile à vérifier en appliquant (5.20) à la fonction positive $|f|$. Notons que $\mathbb{R}^N$ peut être décomposé de plusieurs manière comme $\mathbb{R}^m \times \mathbb{R}^n$ avec les choix différents pour m, n tels que $m + n = N$; évidemment, (5.20) s'applique à toutes ces décompositions. Aussi, par exemple, on peut écrire $\mathbb{R}^3$ comme $\mathbb{R}^1 \times \mathbb{R}^2$ ou $\mathbb{R}^2 \times \mathbb{R}^1$ et ensuite, en écrivant $\mathbb{R}^2$ comme $\mathbb{R}^1 \times \mathbb{R}^1$ on a $3! = 6$ différentes façons équivalentes de calculer $\int_{\mathbb{R}^3} f$ comme le résultat du calcul de trois intégrales unidimensionnelles. Par exemple, en écrivant $x = (x_1, x_2, x_3)$, on a

$$\int_{\mathbb{R}^3} f(x)\,\mathrm{d}x = \int_{\mathbb{R}^1} \left(\int_{\mathbb{R}^1} \left(\int_{\mathbb{R}^1} f(x_1, x_2, x_3)\,\mathrm{d}x_1 \right) \mathrm{d}x_2 \right) \mathrm{d}x_3$$

comme une de ces six possibilités. Nous dirons que l'intégrale triple dans $\mathbb{R}^3$ revient à calculer une des six intégrales itérées équivalentes (pourvu que la fonction à intégrer soit $\geqslant 0$ ou intégrable, sous-entendu mesurable aussi).

La formule (5.20) est appliquée généralement dans les situations où f n'est définie que dans une partie A de $\mathbb{R}^N$; théoriquement, ceci est possible par le prolongement par zéro de f en dehors de A . Illustrons la situation générale en considérant $A \subset \mathbb{R}^2$ (en négligeant toute considération de mesurabilité) ; si $f : A \to \mathbb{R}_+$, écrivons $\widetilde{f}$ pour le prolongement de f par 0 en dehors de A. Alors, si $f \geqslant 0$,

$$\begin{aligned}\iint_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \iint_{\mathbb{R}^2} \widetilde{f}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{-\infty}^{\infty} \mathrm{d}x \left\{ \int_{-\infty}^{\infty} \widetilde{f}(x,y)\,\mathrm{d}y \right\} = \int_{-\infty}^{\infty} \mathrm{d}y \left\{ \int_{-\infty}^{\infty} \widetilde{f}(x,y)\,\mathrm{d}x \right\}.\end{aligned}$$

Si la projection de A sur l'axe des x est un intervalle J dont les extrémités sont $a < b$, alors on aura

$$\iint_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \mathrm{d}x \int_{A_x} f(x,y)\,\mathrm{d}y\,.$$

Dans beaucoup d'applications, A_x sera aussi un intervalle pour chaque $x \in J$; si les extrémités de A_x sont $c(x)$ et $d(x)$ on aura

$$\iint_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \mathrm{d}x \int_{c(x)}^{d(x)} f(x,y)\,\mathrm{d}y\,.$$

Si A_x est la réunion de plusieurs intervalles le principe reste le même, seulement les calculs deviennent plus laborieux ; on écrira

$$\int_{A_x} f(x,y)\,\mathrm{d}y = \sum_i \int_{c_i(x)}^{d_i(x)} f(x,y)\,\mathrm{d}y$$

si A_x est constitué des intervalles d'extrémités $c_i(x)$, $d_i(x)$, $i \geqslant 1$. Du point de vue théorique, A_x peut être sur un ensemble beaucoup plus compliqué mais dans les applications élémentaires, A_x sera généralement un intervalle ou la réunion de deux ou trois intervalles. Cette discussion se généralise au cas de f intégrable au lieu de $f \geqslant 0$. En prenant $f = 1$ sur A on obtient

$$\lambda_2(A) = \text{aire}(A) = \iint_A \mathrm{d}x\,\mathrm{d}y\,.$$

La discussion précédente s'applique également en écrivant

$$\iint_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_\alpha^\beta \mathrm{d}y \int_{A^y} f(x,y)\,\mathrm{d}y$$

où α, β sont les extrémités de la projection de A sur l'axe des y.

Les applications du procédé précédent au cas de $A \subset \mathbb{R}^3$ et $f : A \to \mathbb{R}_+$ ou $\mathbb{R}$ sont analogues, sauf que les possibilités du calcul sont plus variées. Par exemple, s'agissant du calcul de $\lambda_3(A) = \text{volume}\,(A)$ (ici $f \equiv 1$ sur A) on peut procéder comme suit : en écrivant (x,y,z) les éléments de $\mathbb{R}^3$,

$$\lambda_3(A) = \iiint_A \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_a^b \lambda_2(A_z)\,\mathrm{d}z$$

où A_{z_0} est la section de A par le plan $z = z_0$ et les valeurs de z dans A vont de a jusqu'à b ; plus précisément,

$$A_{z_0} = \{(x,y) : (x,\, y,\, z_0) \in A\} \subset \mathbb{R}^2\,.$$

Pour le même problème, on peut aussi procéder comme suit : si B est la projection de A sur le plan (x,y), alors

$$\lambda_3(A) = \iint_B \lambda_1(A_{x,y})\,\mathrm{d}x\,\mathrm{d}y$$

où $A_{x,y}$ est la (x,y)-section de A ; ceci veut dire que

$$A_{x,y} = \{z : (x, y, z) \in A\} \subset \mathbb{R}^1 .$$

Les mêmes procédés appliqués sur les fonctions donnent les formules suivantes.

$$(1) = \iiint_A f(x,y,z)\,dx\,dy\,dz = (2) = \int_a^b dz \left\{ \iint_{A_z} f(x,y,z)\,dx\,dy \right\}$$

$$= (3) = \iint_B dx\,dy \left\{ \int_{A_{x,y}} f(x,y,z)\,dz \right\}$$

où nous avons mis les accolades { } seulement pour faciliter la lecture de ces formules ; noter aussi les écritures des symboles dx, dy, dz ; l'écriture $dx\,dy\,dz$ dans l'intégrale triple de gauche (1) représente l'intégration par rapport à la mesure de λ_3 ; pour la seconde intégrale (2), on pense à λ_3 comme $\lambda_1 \otimes \lambda_2$, dz représentant l'intégrale par rapport à λ_1 et $dx\,dy$ celle par rapport à λ_2 ; pour l'intégrale (3) aussi, on pense à λ_3 comme $\lambda_1 \otimes \lambda_2$ mais cette fois les intégrales sont exécutées dans un ordre différent. Naturellement, le procédé utilisé dans un calcul donné dépend de la commodité que celui-ci peut avoir pour le problème en question. Les exercices du paragraphe 5.4.5 illustrent ce point clairement (cf. Ex. 9-10, § 5.4.5).

Une généralisation de ce qui précède est dans le calcul de $\lambda_N(E)$, $E \subset \mathbb{R}^N$, que nous appellerons quelquefois le ***volume N-dimensionnel*** de E (ce qui est, bien entendu, l'***aire*** si $N = 2$). Nous illustrons cette discussion générale par une série d'exemples donnés dans le paragraphe 5.4.5. Dans ces exercices, la question de mesurabilité ne sera presque jamais soulevée ; tous les ensembles et toutes les fonctions considérés seront supposés mesurables dans le sens approprié.

Une dernière remarque concerne les notations utilisées pour les intégrales par rapport à la mesure de Lebesgue sur les parties de $\mathbb{R}^N$. Si $A \subset \mathbb{R}^N$, $f : A \to \mathbb{R}$, on écrira pour $\int_A f\,d\lambda_N$, selon le besoin, une des expressions suivantes :

$$\int_A f(x)\,dx , \quad \int \cdots \int_A f(x_1, \ldots, x_N)\,dx_1 \cdots dx_N , \quad \int_A f \text{ ou } \int \cdots \int_A f .$$

Si $A = [a, b] \times [c, d] \subset \mathbb{R}^2$, on peut aussi écrire

$$\int_a^b \int_c^d f(x,y)\,dx\,dy \quad \text{ou} \quad \int_c^d \int_a^b f(x,y)\,dx\,dy$$

pour $\int_A f$ ou $\int \int_A f$, la notation étant alors un peu ambigüe car elle ne spécifie plus clairement les domaines de variations respectifs de x et de y. En pratique, toutes ces notations sont utilisées selon les besoins de clarté ou de compacité des formules à traiter.

5.4.5 Exercices

1 Si $A \subset \mathbb{R}^m$, $B \subset \mathbb{R}^n$, $f : A \to \mathbb{R}$, $g : B \to \mathbb{R}$ (où f, g sont intégrables dans A, B respectivement et où f, g sont $\geqslant 0$), alors

$$\int_{A\times B} f(x)g(x)\,dx\,dy = \left(\int_A f(x)\,dx\right)\left(\int_B g(y)\,dy\right).$$

En particulier

$$\int_{A\times B} f(x)f(y)\,dx\,dy = \left(\int_A f(x)\,dx\right)^2.$$

L'exercice suivant donne une série d'exemples illustrant l'utilité de l'exercice 1.

2 Vérifier les formules suivantes.

(i) $\displaystyle\int_0^1\int_0^1 xy\,e^{x^2+y^2}\,dx\,dy = \frac{1}{4}(e-1)^2.$

(ii) $\displaystyle\int_0^1\int_0^1 (x+y)\,e^{x+y}\,dx\,dy = 2(e-1).$

(iii) $\displaystyle\iint_{\substack{0\leqslant x\leqslant 1\\ 1\leqslant y\leqslant 2}} \frac{x}{y}\,dx\,dy = \frac{1}{2}\ln 2.$

(iv) $\displaystyle\int_0^1\int_0^1\int_0^1 e^{(x+y+z)}\,dx\,dy\,dz = (e-1)^3.$

(v) $\displaystyle\iint_{A\times B} \sin(x+y)\,dx\,dy = 4\sin\left(\frac{b-a}{2}\right)\sin\left(\frac{d-c}{2}\right)\sin\left(\frac{a+b+c+d}{2}\right)$

où $A = [a, b]$, $B = [c, d]$ sont les intervalles compacts de $\mathbb{R}$.

(Ecrire $\sin(x+y) = \sin x\cos y + \cos x\sin y$ et utiliser les formules trigonométriques pour $\sin\alpha - \sin\beta$, $\cos\alpha - \cos\beta$.)

(vi) Si $A \subset \mathbb{R}^N$ et $\int_A f^2 < \infty$, $\int_A g^2 < \infty$, alors

$$\frac{1}{2}\iint_{A\times A} \big(f(x)g(y) - f(y)g(x)\big)^2\,dx\,dy = \int_A f^2 \int_A g^2 - \left(\int_A fg\right)^2.$$

(vii) Si $A \subset \mathbb{R}^N$, $\int_A f^2 < \infty$, $\int_A g^2 < \infty$, $\lambda_N(A) < \infty$, alors

$$\frac{1}{2}\iint_{A\times A} \big(f(x) - f(y)\big)\big(g(x) - g(y)\big)\,dx\,dy = \lambda_N(A)\int_A fg - \left(\int_A f\right)\left(\int_A g\right).$$

(Noter que l'inégalité $|ab| \leqslant (a^2 + b^2)/2$ entraîne que fg dans (vi) et (vii) est intégrable sur A.)

3 (Notations de Ex.2(vi)) De l'exercice 2(vi) déduire l'inégalité de Cauchy-Schwarz pour les intégrales :

$$\left|\int_A fg\right| \leqslant \left(\int_A f^2\right)^{\frac{1}{2}} \left(\int_A g^2\right)^{\frac{1}{2}}$$

avec égalité si et seulement si $f(x) = \lambda g(x)$ p.p.(x) dans A pour $\lambda \in \mathbb{R}$ ou $f = 0$ p.p. ou $g = 0$ p.p. dans A.

4 (Notations Ex. 2(vii)) Supposons que $\lambda_N(A) > 0$ et que si x, y sont dans A on a $f(x) - f(y) \geqslant 0$ si et seulement si $g(x) - g(y) \geqslant 0$; alors on a (l'inégalité, appelée aussi de Chebyshev) :

$$\frac{1}{\lambda_n(A)} \int_A fg \geqslant \frac{1}{\lambda_n(A)} \int_A f \cdot \frac{1}{\lambda_n(A)} \int_A g$$

avec égalité si et seulement si f ou g est constante p.p. dans A. Etudier le cas où $f(x) - f(y) \geqslant 0$ si et seulement si $g(x) - g(y) \leqslant 0$.

5 Soit $A \subset \mathbb{R}^2$ un ensemble (mesurable) ayant la propriété que $\lambda_1(A_x) = 0$ p.p.(x) où (x, y) est un élément typique de $\mathbb{R}^2$ et où A_x est la x-section de A ; alors $\lambda_2(A) = 0$. En particulier si A_x est au plus dénombrable p.p.(x), alors $\lambda_2(A) = 0$. Formuler et démontrer les énoncés analogues pour les y-sections de A. En déduire que les ensembles dans $\mathbb{R}^2$ formés par une droite ou un cercle ou une des courbes usuelles sont des ensembles négligeables.

6 Généraliser l'exercice 5 au cas d'une partie A de $\mathbb{R}^3$. En déduire que les ensembles dans $\mathbb{R}^3$ formés par une courbe ou une surface usuelle sont les ensembles négligeables.

7 En suivant l'exercice 6, démontrer qu'un hyperplan quelconque dans $\mathbb{R}^N$ est un ensemble négligeable. Plus généralement, soit $A = \{x : f(x) = 0\}$, $x = (x_1, \ldots, x_N) = (x'; x_N)$, $x' = (x_1, \ldots, x_{N-1})$, et $f : \mathbb{R}^N \to \mathbb{R}$ une fonction borélienne telle que, pour p.p.(x'), $f(x'; x_N) = 0$ possède au plus un nombre dénombrable de solutions x_N (ou, plus généralement, $\lambda_1(A_{x'}) = 0$ p.p.(x')) ; alors A est un ensemble négligeable dans $\mathbb{R}^N$.

(Le cas d'un hyperplan est obtenu en prenant $f(x) = (\alpha \cdot x - a)$, $\alpha \in \mathbb{R}^N$, $a \in \mathbb{R}$.)

8 Représenter les expressions suivantes comme $\int_A$, $A \subset \mathbb{R}^2$ et comme une autre intégrale itérée :

(i) $\displaystyle\int_0^1 \mathrm{d}x \int_x^1 f(x,y)\,\mathrm{d}y$;

(ii) $\displaystyle\int_0^1 \mathrm{d}x \int_{\sqrt{x-x^2}}^{\sqrt{x}} f(x,y)\,\mathrm{d}y$;

(iii) $\displaystyle\int_1^2 \mathrm{d}x \int_x^{2x} f(x,y)\,\mathrm{d}y$;

(iv) $\displaystyle\int_0^\infty \mathrm{d}x \int_x^\infty f(x,y)\,\mathrm{d}y$.

9 Démontrer que

$$\int_0^\infty \mathrm{d}x \int_x^\infty \frac{\varphi(y)}{y}\,\mathrm{d}y = \int_0^\infty \varphi(t)\,\mathrm{d}t$$

si $\varphi : [\,0\,,\,\infty\,[\to \mathbb{R}$ est $\geqslant 0$ ou intégrable.

10 Soit $A = \{(x,y) : 0 \leqslant x \leqslant 1\,,\ 0 \leqslant y \leqslant x^2\}$; si $f(x,y) = 1/(1+x^2)$ calculer $\int_A f$ de deux manières différentes.

11 Calculer les intégrales suivantes (dans $\mathbb{R}^2$) :

(i) $\displaystyle\iint_A \mathrm{e}^{-(x+y)}\,\mathrm{d}x\,\mathrm{d}y$ où $A = D \times D$,
D un ensemble au plus dénombrable dans $\mathbb{R}$;

(ii) $\displaystyle\iint_A \sin(xy)\,\mathrm{d}x\,\mathrm{d}y$ où $A = \bigcap_{n=1}^{\infty} A_n$, $A_n = [\,0\,,\,1\,] \times [\,0\,,\,1/n\,]$;

(iii) $\displaystyle\iint_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$ où $f = \prod_{n=1}^{\infty} 1_{A_n}$, A_n étant comme dans (ii).

(Toutes ces intégrales valent 0.)

12 Calculer

$$\lim_{n\to\infty} \iint_A x^{1/n} y^{1/n} \sin\left(x + y + \frac{1}{n}\right) \mathrm{d}x\,\mathrm{d}y$$

où $A = [\,0\,,\,\pi/2\,] \times [\,0\,,\,\pi/2\,]$.

13 Vérifier que si $f : \mathbb{R} \to \mathbb{R}$ est intégrable ou $\geqslant 0$, alors

$$\left(\int_{\mathbb{R}} f\right)^2 = 2\int_{-\infty}^{\infty} \mathrm{d}x \int_x^\infty f(x) f(y)\,\mathrm{d}y\,.$$

14 Démontrer que

$$\iint_{A_r} (x^2+y^2)\,\mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y \leqslant \frac{\pi r^2}{\mathrm{e}}, \quad r>0,$$

où $A_r = \{(x,y) : x^2+y^2 \leqslant r^2\} \subset \mathbb{R}^2$.

15 Les intégrales suivantes représentent $\lambda_2(A)$ (l'aire de A) pour certains $A \subset \mathbb{R}^2$; dessiner ces ensembles A et déterminer $\lambda_2(A)$ dans chaque cas :

(i) $\displaystyle\int_0^1 \mathrm{d}y \int_{y-1}^{-y+1} \mathrm{d}x + \int_{-1}^0 \mathrm{d}y \int_{-y-1}^{y+1} \mathrm{d}x$;

(ii) $\displaystyle\int_0^1 \mathrm{d}x \int_{x^2}^{\sqrt{x}} \mathrm{d}y$;

(iii) $\displaystyle\int_{-a}^{a} \mathrm{d}x \int_{-b\sqrt{1-(x/a)^2}}^{b\sqrt{1-(x/a)^2}} \mathrm{d}y$.

16 Si $A \subset \mathbb{R}^3$ on a vu que :

$$\text{(a)}\ \ \lambda_3(A) = \int_a^b \lambda_2(A_z)\,\mathrm{d}z\,; \qquad \text{(b)}\ \ \lambda_3(A) = \iint_B \lambda_1(A_{x,y})\,\mathrm{d}x\,\mathrm{d}y.$$

Calculer les volumes des ensembles suivants par la méthode suggérée dans chaque cas :

(i) $A = \left\{(x,y,z) : \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 \leqslant 1\right\}$;
méthode (a) : $(a, b, c > 0)$;

(ii) $A = \left\{(x,y,z) : x \geqslant 0\,,\ y \geqslant 0\,,\ z \geqslant 0\,,\ \frac{x}{a} + \frac{y}{b} + \frac{z}{c} \leqslant 1\right\}$;
méthode (b) : $(a,\ bc > 0)$;

(iii) $A = \{(x,y,z) : x^2+y^2 \leqslant z \leqslant x\}$; méthode (a) ;

(iv) $A = \{(x,y,z) : x^2+y^2 \leqslant a^2\,,\ x^2+z^2 \leqslant a^2\}$;
méthode (a), en travaillant avec les x-sections $(a > 0)$.

Décrire géométriquement chacun des ensembles A en question.

17 Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue ; pour $a > 0$ posons $D_a = [-a\,,\ a] \times [-a\,,\ a]$. Démontrer que

$$\iint_{D_a} f(x-y)\,\mathrm{d}x\,\mathrm{d}y = \begin{cases} 0 & \text{si } f \text{ est impaire} \\ 2\displaystyle\int_0^{2a} (2a-t)f(t)\,\mathrm{d}t & \text{si } f \text{ est paire} \end{cases}$$

(f impaire $\Leftrightarrow f(-x) = -f(x)$, $x \in \mathbb{R}$; f paire $\Leftrightarrow f(-x) = f(x)$, $x \in \mathbb{R}$).

Les exercices suivants montrent comment on peut calculer certaines intégrales dans $\mathbb{R}$ en passant aux intégrales doubles.

18 Soit $f :]0, \infty[\to \mathbb{R}$ une fonction continue intégrable ; posons $\alpha = \int_0^\infty f$ et F une primitive de f. On va démontrer que, pour $0 < a < b$,

$$I = \int_0^\infty \frac{F(bt) - F(at)}{t}\, \mathrm{d}t = \alpha \ln\Big(\frac{b}{a}\Big)$$

(formule de Cauchy-Frullani). Vérifier d'abord que

$$\int_0^\infty \left|\frac{F(bt) - F(at)}{t}\right| \mathrm{d}t \leqslant \beta(b-a) < \infty$$

où $\beta = \int_0^\infty |f|$ (en utilisant le théorème de Fubini). Donc, toujours par le théorème de Fubini,

$$I = \int_0^\infty \frac{\mathrm{d}t}{t} \int_{at}^{bt} f(u)\, \mathrm{d}u = \int_0^\infty \mathrm{d}t \int_a^b f(xt)\, \mathrm{d}x$$
$$= \int_a^b \mathrm{d}x \int_0^\infty f(xt)\, \mathrm{d}t = \alpha \int_a^b \frac{\mathrm{d}x}{x} = \alpha \ln\Big(\frac{b}{a}\Big).$$

19 De l'exercice 18 déduire que, si $0 < a < b$,

$$\int_0^\infty \frac{\mathrm{e}^{-at} - \mathrm{e}^{-bt}}{t}\, \mathrm{d}t = \ln\Big(\frac{b}{a}\Big)$$
$$\int_0^\infty \frac{\operatorname{arc\,tg}(bt) - \operatorname{arc\,tg}(at)}{t}\, \mathrm{d}t = \frac{\pi}{2} \ln\Big(\frac{b}{a}\Big).$$

20 Etablir que

$$I = \lim_{T\to\infty} \int_0^T \frac{\sin x}{\sqrt{x}}\, \mathrm{d}x = \int_0^\infty \frac{\sin x}{\sqrt{x}}\, \mathrm{d}x = \sqrt{\frac{\pi}{2}}$$
$$J = \lim_{T\to\infty} \int_0^T \sin x^2\, \mathrm{d}x = \int_0^\infty \sin x^2\, \mathrm{d}x = \frac{1}{2}\sqrt{\frac{\pi}{2}}$$

(intégrales impropres) en justifiant le raisonnement suivant :

$$\int_0^\infty \mathrm{e}^{-t^2 x}\, \mathrm{d}t = \frac{1}{2}\sqrt{\frac{\pi}{x}}, \quad x > 0 \text{ (Ex. 1, § 5.3.5)}$$
$$\frac{2}{\sqrt{\pi}} \int_0^\infty \mathrm{e}^{-t^2 x} \sin x\, \mathrm{d}t = \frac{\sin x}{\sqrt{x}}$$
$$\int_0^T \frac{\sin x}{\sqrt{x}}\, \mathrm{d}x = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{-\cos T - t^2 \sin T}{1+t^4} \mathrm{e}^{-Tt^2}\, \mathrm{d}t + \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\mathrm{d}t}{1+t^4}$$

d'où

$$I = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\mathrm{d}t}{1+t^4}.$$

On verra en analyse complexe que $\int_0^\infty \mathrm{d}t/(1+t^4) = \pi/(2\sqrt{2})$.

Les exercices suivants donnent quelques contre-exemples associés au théorème de Fubini.

21 Soit u une fonction réelle de classe C^2 définie dans un ensemble ouvert contenant $[0,1]\times[0,1]\setminus\{(0,0)\}$ telle que

$$\lim_{x\to 0+} u(x,0)=\alpha\,,\qquad \lim_{y\to 0+} u(0,y)=\beta$$

où α, β sont dans $\mathbb{R}$; soit $f=D_{12}u$ sur $]0,1[\times]0,1[$. Vérifier que

$$\int_0^1 \mathrm{d}x\int_0^1 f(x,y)\,\mathrm{d}y=\alpha+k\,,\qquad \int_0^1 \mathrm{d}y\int_0^1 f(x,y)\,\mathrm{d}x=\beta+k$$

où $k=u(1,1)-u(0,1)-u(1,0)$. Comme cas particulier on a les exemples suivants où les deux intégrales itérées donnent des résultats différents :

(i) $u(x,y)=\dfrac{x}{x+y}\,,\quad f(x,y)=\dfrac{x-y}{(x+y)^3}$;

(ii) $u(x,y)=\theta(x,y)$ (angle polaire), $\quad f(x,y)=\dfrac{y^2-x^2}{\left(x^2+y^2\right)^2}$;

(iii) $u(x,y)=\dfrac{x^2-y^2}{x^2+y^2}\,,\quad f(x,y)=8xy\dfrac{x^2-y^2}{\left(x^2+y^2\right)^3}$.

Naturellement, $|f|$ n'est pas intégrable dans les exemples (i)-(iii).

22 Soit $f(x,y)=\exp(-xy)-2\exp(-2xy)$; vérifier que

$$\int_0^1 \mathrm{d}y\int_1^\infty f(x,y)\,\mathrm{d}x>0\,,\qquad \int_1^\infty \mathrm{d}x\int_0^1 f(x,y)\,\mathrm{d}y<0\,.$$

Conclure que $\int_D |f|=\infty$, où $D=[1,\infty[\times[0,1]$.

23 Soit a une fonction réelle de classe C^2 définie dans $[0,\infty[$ telle que $\lim_{t\to\infty} a(t)$ et $\lim_{t\to\infty} ta'(t)$ existent (dans $\mathbb{R}$) ; soient $u(x,y)=a(y/x)$, $f(x,y)=D_{12}u(x,y)$. Vérifier que

$$\int_1^\infty \mathrm{d}x\int_1^\infty f(x,y)\,\mathrm{d}y=a(1)-a(0)\,,\qquad \int_1^\infty \mathrm{d}y\int_1^\infty f(x,y)\,\mathrm{d}x=a(1)-a(\infty)\,.$$

Vérifier que si $\lim_{t\to\infty} ta''(t)$ existe dans $\mathbb{R}$, alors f est une fonction bornée dans $[1,\infty[\times[1,\infty[$. Vérifier que les choix suivants de a donnent des exemples où les deux intégrales itérées sont différentes même si les f concernées sont des fonctions rationnelles, continues et bornées dans $[1,\infty[\times[1,\infty[$:

(i) $a(t)=(1+t)^{-1}$; $\quad f(x,y)=(x-y)(x+y)^{-3}$;

(ii) $a(t)=\operatorname{arc\,tg} t$; $\quad f(x,y)=(y^2-x^2)(x^2+y^2)^{-2}$.

24 Dans l'exercice 23, prendre $a(t) = t/(1+t^2)$; alors on aura

$$f(x,y) = (6x^2y^2 - x^4 - y^4)(x^2+y^2)^{-3} ;$$

vérifier que, dans ce cas, les deux intégrales itérées existent et sont égales à $1/2$ mais que $\int_D |f| = \infty$ où $D = [1, \infty[\times [1, \infty[$. (Pour le dernier point, voir que $f(x,y)$ est de la forme $x^{-2}g(y/x)$, g continue sur $[0, \infty[$; vérifier que $\int_D |f| = \infty$ pour de telles fonctions f.)

25 Soit $f(x,y) = xy(x^2+y^2+1)^{-2}$, $(x,y) \in \mathbb{R}^2$. Vérifier que $\int_{\mathbb{R}^2} |f| = \infty$ mais que chacune des deux intégrales itérées existe et est égale à 0.

5.5 MESURES DANS $\mathbb{R}^N$

La théorie développée dans les sections 5.2 à 5.4 est très générale ; elle concerne l'intégration dans un espace mesuré quelconque, même si les exemples, les exercices et la motivation sous-jacente au développement sont tous basés sur l'intégrale de Lebesgue dans $\mathbb{R}^N$. Il est important d'observer de nouveau que le développement de cette théorie générale est motivé, pour nous, par le désir d'avoir une théorie qui soit facile d'application dans $\mathbb{R}^N$. Mais, il est intéressant de remarquer aussi, que la théorie générale fonctionne très efficacement sur la seule base de l'additivité dénombrable de la mesure positive (supposee σ-finie dans la section 5.4) ; la théorie des sections 5.2 à 5.4 ne deviendrait pas plus simple si nous avions décidé de nous restreindre au cas de la mesure de Lebesgue dans $\mathbb{R}^N$.

Dans cette section, nous nous bornons au cas des mesures dans $\mathbb{R}^N$ pour préparer le terrain pour la démonstration d'un théorème important du sujet : le théorème du changement de variables dans les intégrales multiples ; le théorème lui-même sera le sujet de la section 5.7.

5.5.1 Mesures de Radon

Soient Ω une partie ouverte de $\mathbb{R}^N$, Σ la tribu borélienne de Ω ; rappelons que Σ est la plus petite tribu contenant toutes les parties ouvertes contenues dans Ω. Nous dirons que $\mu : \Sigma \to [0, \infty]$ est une ***mesure de Radon*** (***positive***) dans Ω si μ est une mesure positive telle que $\mu(K) < \infty$ pour toute partie compacte K contenue dans Ω. Une mesure de Radon positive μ dans Ω (Ω étant ouvert dans $\mathbb{R}^N$) possède automatiquement les propriétés suivantes :

- *μ est σ-finie ;*
- *regularité intérieure de μ : si $A \in \Sigma$,*

$$\mu(A) = \sup\{\mu(K) : K \subset A,\ K \text{ compact}\} ;$$

- *régularité extérieure de μ : si $A \in \Sigma$,*

$$\mu(A) = \sup\{\mu(V) : V \supset A,\ V \text{ ouvert}\} .$$

Le fait que μ est σ-finie est immédiat car Ω est la réunion d'une suite d'ensembles compacts. La démonstration de la régularité de μ est plus longue et nous ne la ferons pas ici (cf. la section 5.13 pour d'autres remarques à ce sujet et au sujet de la terminologie adoptée ici).

Les propriétés de régularité de μ ont comme conséquence importante et immédiate que μ est uniquement déterminée par ces valeurs sur les ensembles compacts ou ouverts ; ceci est formulé précisément et plus généralement par l'énoncé suivant.

Unicité (I)

Soient μ, ν deux mesures de Radon positives dans Ω où Ω est un ensemble ouvert dans $\mathbb{R}^N$. Si pour toute partie compacte $K \subset \Omega$, $\mu(K) \leqslant \nu(K)$ alors $\mu(A) \leqslant \nu(A)$, pour toute partie borélienne $A \subset \Omega$. Si pour toute partie ouverte $V \subset \Omega$, $\mu(V) \leqslant \nu(V)$ alors $\mu(A) \leqslant \nu(A)$, pour toute partie borélienne $A \subset \Omega$.

En fait, il y a un énoncé d'unicité qui est plus général. Pour le donner, rappelons qu'un ***cube fermé*** dans $\mathbb{R}^N$ est un pavé compact dont les côtés sont de longueurs égales ; ainsi, un cube fermé est un ensemble C de la forme $I_1 \times \cdots \times I_N$ où chaque I_j est un intervalle compact avec $\lambda_1(I_1) = \ldots = \lambda_1(I_N)$. Nous dirons que C est un ***cube fermé à gauche*** (resp. ***fermé à droite***) si $C = I_1 \times \cdots \times I_N$ où chaque I_j est un intervalle borné de la forme $[\,a_j\,,\,b_j\,[$ (resp. $]\,a_j\,,\,b_j\,]$) avec $b_1 - a_1 = \ldots = b_N - a_N$ (c'est-à-dire $\lambda_1(I_1) = \ldots = \lambda_1(I_N)$). La définition d'un ***pavé borné fermé à gauche*** (resp. ***fermé à droite***) est similaire sauf que l'on ne demande pas l'égalité des longueurs des côtés. Nous énonçons maintenant le second théorème d'unicité.

Unicité (II)

Soient μ, ν comme ci-dessus (unicité (I)). Si pour tout cube fermé (ou tout cube fermé à gauche ou tout cube fermé à droite) $C \subset \Omega$ on a $\mu(C) \leqslant \nu(C)$, alors $\mu(A) \leqslant \nu(A)$ pour toute partie borélienne $A \subset \Omega$.

Le passage des cubes fermés aux cubes fermés à gauche ou à droite et vice versa est immédiat ; en effet, tout cube fermé C est la limite monotone (décroissante) des cubes fermés à gauche où à droite et vice-versa (avec la limite monotone croissante). La proposition 5.5.2 et la remarque 5.5.3 suivantes donnent, entre autres choses, qu'une partie ouverte de $\mathbb{R}^N$ est une réunion dénombrable des cubes fermés à gauche, deux à deux disjoints ; ceci permet d'appliquer le théorème d'unicité (I) et de compléter la démonstration d'unicité (II).

Pour conclure ce paragraphe, donnons quelques exemples importants de mesure de Radon dans $\mathbb{R}^N$ ou dans une partie ouverte Ω de $\mathbb{R}^N$. L'exemple le plus important est sans doute la mesure de Lebesgue λ_N dans $\mathbb{R}^N$ ou sa restiction à Ω. Une classe importante d'exemples est donnée par des densités par rapport à λ_N :

$$\mu(A) = \int_A p \, \mathrm{d}\lambda_N \,, \quad A \subset \Omega \,, \; A \text{ borélien} \,,$$

avec $p : \Omega \to [0, \infty[$, une fonction borélienne qui est intégrable sur chaque ensemble compact $K \subset \Omega$ (cf. Ex. 25-27, § 5.2.15). Une autre famille d'exemples est fournie par la construction indiquée par l'exercice 2 (§ 5.2.15) ; on choisit un ensemble au plus dénombrable $X = \{x_1, x_2, \ldots\} \subset \mathbb{R}^N$ et une suite de nombres $p_i \geqslant 0$, $i \geqslant 1$, tels que $\sum_i p_i < \infty$; on pose alors

$$\mu(A) = \sum_{i=1}^{\infty} p_i \, 1_A(x_i) \,, \quad A \subset \mathbb{R}^N \,.$$

Ceci donne une mesure positive finie dans $\mathbb{R}^N$ définie pour tous les sous-ensembles de $\mathbb{R}^N$. Si $\sum_i p_i = \infty$, il faut imposer une condition sur X pour que $\mu(K) < \infty$ pour tout K compact ; il suffirait, par exemple, de supposer que X soit un ensemble qui n'a aucun point d'accumulation dans $\mathbb{R}^N$ de sorte que $K \cap X$ est un ensemble de cardinalité finie quel que soit l'ensemble compact K. Les mesures positives de ce type s'appellent les mesures positives ***purement atomiques***. On peut imaginer des mesures de Radon formées par des combinaisons linéaires de mesures ayant une densité par rapport à λ_N et une autre qui est purement atomique. Il existe des mesures de Radon positives plus compliquées encore qui sont « portées » par des ensembles de « dimension » k, $0 \leqslant k \leqslant N$, un exemple simple d'un tel ensemble étant une sous-variété différentielle de $\mathbb{R}^N$ de dimension k.

La proposition suivante n'a pas une relation directe avec la théorie de l'intégration ; néanmoins, elle est très utile pour différentes considérations analytiques, y comprise la théorie de l'intégration.

5.5.2 Proposition

Soit V un ensemble ouvert non vide dans $\mathbb{R}^N$. Pour tout $\delta > 0$, on peut représenter V comme la réunion d'une suite de cubes fermés $C_1, C_2, \ldots$ tels que :

(i) les côtés des C_i sont de longueur $\delta 2^{-n}$ pour $n = 0, 1, 2, \ldots$,
(ii) $\mathring{C}_i \cap \mathring{C}_j = \emptyset$ si $i \neq j$.

On dira que les cubes $\{C_i\}$ sont deux à deux ***non empiétants*** s'ils satisfont la condition (ii).

DÉMONSTRATION

Observons d'abord que $\mathbb{R}^N$ est la réunion des cubes fermés, deux à deux non empiétants, $\{A_n^{(1)}\}$, $n = 1, 2, \ldots$, ayant des côtés de longueurs δ. Soient $\{C_n^{(1)}\}$ les cubes parmi les $\{A_n^{(1)}\}$ qui sont contenus dans V et rangeons les autres cubes du système $\{A_n^{(1)}\}$ comme $B_1^{(1)}, B_2^{(1)}, \ldots$ Représentons chaque cube $B_n^{(1)}$ comme une réunion de cubes fermés, deux à deux non empiétants, ayant des côtés de longueurs $\delta/2$ et rangeons tous les cubes ainsi obtenus comme $\{A_n^{(2)}\}$. Soient $\{C_n^{(2)}\}$ les cubes parmi les $\{A_n^{(2)}\}$ qui sont contenus dans V et $\{B_n^{(2)}\}$ les autres cubes parmi les $\{A_n^{(2)}\}$. En continuant ce procédé, on aura les cubes fermés

$$\{A_n^{(m)}\}, \{C_n^{(m)}\}, \{B_n^{(m)}\}, \quad n, m = 1, 2, \ldots$$

Les cubes $C_n^{(m)}$ sont contenus dans V et ils sont deux à deux non empiétants ; de plus, les longueurs des côtés de $C_n^{(m)}$ sont $\delta 2^{-(m-1)}$; démontrons que leur réunion est V. En effet, soit $x \in V$; comme V est ouvert, il existe un $\varepsilon > 0$ tel que le cube fermé $C(x;\varepsilon)$, centré à x de côtés de longueur 2ε, est contenu dans V ; rappelons que

$$C(x;\varepsilon) = \{y \in \mathbb{R}^N : |y - x| \leqslant \varepsilon\}.$$

Si $0 < r < \varepsilon/2$ et $x \in C(b;r)$ alors $C(b;r) \subset C(x;\varepsilon) \subset V$, ce qui résulte de l'inégalité

$$|y - x| \leqslant |y - b| + |x - b| \leqslant r + r < \varepsilon$$

valable si $y \in C(b;r)$.

Soit m un entier $\geqslant 1$ tel que $\delta/2^m < \varepsilon/2$; si $x \notin C_n^{(j)}$, $j = 1, 2, \ldots, (m-1)$, $n \geqslant 1$, alors $x \in A_\alpha^{(m)}$ pour un indice α ; mais $A_\alpha^{(m)}$ est un cube fermé du type $C(b;r)$, $r = \delta/2^m < \varepsilon/2$, avec $x \in C(b;r)$ de sorte que, selon ce qui précède,

$$x \in A_\alpha^{(m)} \subset C(x;\varepsilon) \subset V.$$

Ainsi, $A_\alpha^{(m)}$ est un des cubes $\{C_n^{(m)}\}$ ce qui démontre que V est la réunion des cubes $C_n^{(m)}$, $n \geqslant 1$, $m \geqslant 1$. En rangeant les cubes $C_n^{(m)}$ dans une suite $C_1, C_2, \ldots$ on achève la démonstration. ♦

5.5.3 Remarque

Dans la proposition 5.5.2 on peut choisir les cubes C_n tel qu'ils soient tous fermés à gauche (ou tous fermés à droite) et deux à deux disjoints. Il suffit de répéter la démonstration précédente en remplaçant les mots « cubes fermés » partout par « cubes fermés à gauche » (ou « cubes fermés à droite ») et les mots « non empiétants » par les mots « deux à deux disjoints ».

5.6 MESURE DE LEBESGUE DANS $\mathbb{R}^N$

5.6.1 Généralités

Dans cette section nous spécialisons les considérations de la section précédente au cas de la mesure de Lebesgue dans $\mathbb{R}^N$. Nous caractérisons celle-ci comme l'unique mesure positive de Radon dans $\mathbb{R}^N$ (à une constante multiplicative près) qui est invariante par translation. Nous étudions ensuite le comportement de la mesure de Lebesgue sous des transformations de l'espace sous-jacent $\mathbb{R}^N$, d'abord sous des transformations affines et ensuite sous des transformations non linéaires de classe C^1. Ceci permet d'établir le théorème du changement des variables dans les intégrales multiples (sect. 5.7) et de démontrer la négligeabilité dans $\mathbb{R}^N$ des sous-variétés différentielles de dimension $k < N$.

Si $A \subset \mathbb{R}^N$, $B \subset \mathbb{R}^N$, $a \in \mathbb{R}^N$, $\lambda \in \mathbb{R}^*$, nous posons

$$A + B = \{z \in \mathbb{R}^N : z = x + y,\ x \in A,\ y \in B\},$$
$$A + a = \{z \in \mathbb{R}^N : z = x + a,\ x \in A\},$$
$$\lambda A = \{z \in \mathbb{R}^N : z = \lambda x,\ x \in A\}.$$

Si A est borélien, λA et $A + a$ sont boréliens ; en général, nous ne démontrons pas ce genre d'affirmations, poursuivant en cela notre pratique de ne pas insister sur les questions de mesurabilité. En l'occurence, ces démonstrations sont faciles ; si $\lambda \neq 0$, $f(x) = \lambda^{-1}x$ est une fonction continue, donc borélienne d'où $f^{-1}(A) = \lambda A$ est borélien si A est borélien ; de même, la considération de $f(x) = x - a$ avec $f^{-1}(A) = A + a$ donne le résultat sur $A + a$. Lorsque la question de mesurabilité devient vraiment subtile, nous la signalerons sans nécessairement donner toutes les démonstrations exigées. Par exemple, $A + B$ n'est pas nécessairement borélien, même si A, B le sont ; on peut démontrer, par contre, que $A + B$ est λ_N-mesurable si A, B sont deux parties boréliennes dans $\mathbb{R}^N$. Pour la suite, il suffira de savoir que $A + B$ est ouvert, donc borélien, si A ou B est ouvert ; en effet, si B est ouvert, la formule

$$A + B = \bigcup_{a \in A} \{B + a\}$$

présente $A + B$ comme une réunion des ensembles ouverts $B + a$, $a \in A$, d'où $A + B$ est ouvert ; on peut démontrer facilement que $A + B$ est compact si chacun de A et B est compact. Nous n'aurons pas besoin d'autres considérations de mesurabilité sur $A + B$.

Une mesure positive μ définie sur la tribu borélienne de $\mathbb{R}^N$ s'appelle ***invariante par translation*** si

$$\mu(A + a) = \mu(A), \quad a \in \mathbb{R}^N,\ A \subset \mathbb{R}^N,\ A \text{ borélien}.$$

5.6.2 Proposition

Soit μ une mesure de Radon positive, invariante par translation, dans $\mathbb{R}^N$. Il existe alors un nombre $t \geqslant 0$ tel que $\mu = t\lambda_N$. La mesure λ_N est caractérisée par le fait que c'est une mesure de Radon positive, invariante par translation, dans $\mathbb{R}^N$ avec $\lambda_N(E) = 1$ où E est le cube fermé $[0, 1]^N$.

DÉMONSTRATION

Démontrons d'abord que λ_N est invariante par translation. Nous supposerons dans la suite que tous les ensembles sont boréliens. Pour $a \in \mathbb{R}^N$, posons

$$\nu_a(A) = \lambda_N(A + a), \quad A \subset \mathbb{R}^N .$$

On vérifie immédiatement que ν_a est une mesure de Radon positive telle que $\nu_a(C) = \lambda_N(C)$ pour tout cube fermé C ; le théorème d'unicité (II) (§ 5.5.1) établit alors que $\nu_a = \lambda_N$; donc λ_N est invariante par translation.

Soit μ une mesure de Radon positive quelconque dans $\mathbb{R}^N$, invariante par translation ; soit $\mu(F) = t$ où $F = [0, 1[^N$, un cube fermé à gauche de côtés de longueur 1. Si $t = 0$, on aura $\mu(\mathbb{R}^N) = 0$ car $\mathbb{R}^N$ est la réunion dénombrable des cubes $F_1, F_2, \ldots$, fermés à gauche, deux à deux disjoints, chacun étant un translaté de F d'où

$$\mu(\mathbb{R}^N) = \sum_{i=1}^{\infty} \mu(F_i) = \sum_{i=1}^{\infty} \mu(F) = 0 .$$

Si $t > 0$, posons $\nu = (1/t)\mu$; démontrons que $\nu(C) = \lambda_N(C)$ pour tout cube C fermé à gauche de côtés de longueur 2^{-n}, $n = 0, 1, 2, \ldots$ Comme $\nu(F) = \lambda_N(F) = 1$ pour $F = [0, 1[^N$, ceci est vrai pour $n = 0$ car tout cube C fermé à gauche de côté de longueur 1 est un translaté de F. Comme F est la réunion de 2^{nN} cubes F_i fermés à gauche ayant des côtés de longueur 2^{-n} qui sont deux à deux disjoints et qui sont chacun un translaté des autres, on aura que

$$\nu(F_i) = 2^{-nN}\nu(F) = 2^{-nN} = \lambda_N(F_i)$$

d'où $\nu(C) = \lambda_N(C)$ pour tout cube C fermé à gauche de côtés de longueur 2^{-n}.

Selon la remarque 5.5.3, tout ensemble ouvert de $\mathbb{R}^N$ est la réunion dénombrable des cubes fermés à gauche, deux à deux disjoints, ayant des côtés de longueurs 1 ou 1/2 ou $1/2^2$, etc. ; on en conclut que $\nu(V) = \lambda_N(V)$ pour tout ensemble ouvert $V \subset \mathbb{R}^N$ d'où, selon le théorème d'unicité (I) (§ 5.5.1), $\nu = \lambda_N$. ♦

5.6.3 Transformations affines

Une application $\varphi : \mathbb{R}^N \to \mathbb{R}^N$ telle que

$$\varphi(x) \doteq Tx + a, \quad x \in \mathbb{R}^N ,$$

où $a \in \mathbb{R}^N$ et $T : \mathbb{R}^N \to \mathbb{R}^N$ est une application linéaire, s'appelle une application affine.

Nous aurons besoin du résultat suivant sur une application linéaire *inversible* T : *il existe deux matrices orthogonales* P, Q *et une matrice diagonale* $D = \operatorname{diag}[t_1, \ldots, t_N]$ *avec les* $t_i > 0$ *telles que*

$$T = PDQ \quad \text{avec } |\det T| = t_1 \cdots t_N . \tag{5.21}$$

Ici, comme ailleurs dans ce livre, nous avons identifié une transformation linéaire de $\mathbb{R}^N$ dans $\mathbb{R}^N$ avec sa matrice canonique ; rappelons aussi qu'une matrice S du type (N, N) s'appelle une matrice orthogonale si

$${}^tSS = S\,{}^tS = I$$

où I est la matrice identité du type (N, N). Pour la démonstration de (5.21), il suffit de diagonaliser la matrice symétrique de type strictement positif ${}^tTT$; il existe une matrice orthogonale S et une matrice diagonale $D_1 = \operatorname{diag}[d_1, \ldots, d_N]$, les d_i étant > 0, telles que

$${}^tS({}^tTT)S = D_1 .$$

Si $t_i = \sqrt{d_i}$, $i = 1, \ldots, N$, et $D = \operatorname{diag}[t_1, \ldots, t_N]$ on voit que

$$D_1 = D \cdot D$$

d'où l'on déduit que

$$D^{-1}\,{}^tS({}^tTT)SD^{-1} = I$$

ce qui montre que $P = TSD^{-1}$ est une matrice orthogonale. On remarque alors immédiatement que

$$T = PDS^{-1} ,$$

ce qui donne (5.21) avec $Q = S^{-1}$. Comme le déterminant d'une matrice orthogonale est ± 1, on obtient que

$$|\det T| = t_1 \cdots t_N = \det D . \qquad \blacklozenge$$

Notons une propriété caractéristique d'une matrice orthogonale S : pour tout $x \in \mathbb{R}^N$, $\|Sx\| = \|x\|$; ceci entraîne que si B est une boule de centre 0 dans $\mathbb{R}^N$, alors $S(B) = B$ de sorte que $\lambda_N\big(S(B)\big) = \lambda_N(B)$.

5.6.4 Proposition

Soit $\varphi : \mathbb{R}^N \to \mathbb{R}^N$ *une transformation affine donnée par* $\varphi(x) = Tx + a$, $x \in \mathbb{R}^N$. *Alors*

$$\lambda_N\big(\varphi(A)\big) = |\det T|\,\lambda_N(A) , \quad A \subset \mathbb{R}^N , \ A \textit{ borélien.}$$

DÉMONSTRATION

Comme $\varphi(A) = T(A) + a$ et λ_N est invariante par translation, on a

$$\lambda_N\big(\varphi(A)\big) = \lambda_N\big(T(A)\big) .$$

Supposons d'abord que T est inversible et posons

$$\mu(A) = \lambda_N\big(T(A)\big)\,, \quad A \subset \mathbb{R}^N\,.$$

De la linéarité de T et de l'invariance de λ_N par translation, on déduit que μ est une mesure de Radon positive, invariante par translation ; de la proposition 5.6.2, on conclut qu'il existe $t \geqslant 0$ tel que

$$\lambda_N\big(T(A)\big) = t\,\lambda_N(A)\,, \quad A \subset \mathbb{R}^N\,. \tag{5.22}$$

Il reste à voir que $t = |\det T|$; ceci est vrai si T est une matrice diagonale $D = \operatorname{diag}[t_1, \ldots, t_N]$ avec les $t_i > 0$ car alors

$$T\big([0,1] \times \cdots \times [0,1]\big) = [0, t_1] \times \cdots \times [0, t_N]\,;$$

d'où en prenant $A = [0,1]^N$, $D = T$ dans (5.22), on conclut que $t = t_1 \cdots t_N$. En prenant T comme une matrice orthogonale et B une boule fermée de centre 0 et rayon 1 dans (5.22), on voit que dans ce cas $t = 1$ car alors $T(B) = B$. Passons maintenant au cas général en utilisant (5.21) et $A = B$ dans (5.22) ; on aura

$$\lambda_N\big(T(A)\big) = \lambda_N\big(P\big(DQ(B)\big)\big) = \lambda_N\big(D\big(Q(B)\big)\big) = t\lambda_N\big(Q(B)\big) = t\lambda_N(B)$$

avec $t = \det D = |\det T|$.

Prenons maintenant le cas d'une transformation linéaire T non inversible ; alors $T(\mathbb{R}^N)$ est un sous-espace linéaire de dimension $k < N$, $k = \operatorname{rang} T$ ce qui entraîne que $\lambda_N\big(T(\mathbb{R}^N)\big) = 0$ (cf. Ex. 7, § 5.4.5 ; noter que chaque sous-espace linéaire de dimension $k < N$ est contenu dans un sous-espace linéaire de dimension $(N-1)$). Ainsi

$$\lambda_N\big(T(A)\big) = 0 \quad \text{pour } A \subset \mathbb{R}^N$$

car $T(A) \subset T(\mathbb{R}^N)$. ♦

5.6.5 Volume d'un parallélépipède dans $\mathbb{R}^N$

Soient $\xi_1, \ldots, \xi_N$, N vecteurs dans $\mathbb{R}^N$; posons, pour $d = 1, \ldots, N$,

$$P[\xi_1, \ldots, \xi_d] = \left\{ x \in \mathbb{R}^N : x = \sum_{j=1}^{d} t_j \xi_j\,,\ 0 \leqslant t_j \leqslant 1\,,\ j = 1, \ldots, d \right\}.$$

Nous appellerons $P[\xi_1, \ldots, \xi_d]$ le ***parallélépipède d-dimensionnel*** dans $\mathbb{R}^N$ engendré par les vecteurs $\xi_1, ., ., \xi_d$; c'est le plus petit ensemble convexe contenant $\xi_1, \ldots, \xi_d$; pour $d = 1$ on obtient le segment $[0, \xi_1] \subset \mathbb{R}^N$; pour $d = N$, nous appellerons $P[\xi_1, \ldots, \xi_N]$ le ***parallélépipède*** dans $\mathbb{R}^N$ engendré par $\xi_1, \ldots, \xi_N$. Notons que si $\xi_i = e_i$, $i = 1, \ldots, N$ (les e_i forment la

base canonique de $\mathbb{R}^N$), alors $P[e_1, \ldots, e_N] = [0,1]^N$. La proposition 5.6.4 permet de calculer le volume N-dimensionnel de $P[\xi_1, \ldots, \xi_N]$: *si*

$$P = P[\xi_1, \ldots, \xi_N], \quad \textit{alors } \lambda_N(P) = |\det T|, \tag{5.23}$$

où T est la matrice de type (N,N) formée des N vecteurs colonnes $\xi_1, \ldots, \xi_N$ dans $\mathbb{R}^N$.

La démonstration de (5.23) consiste dans la remarque que, si T est la matrice indiquée, on a

$$T([0,1]^N) = P,$$

d'où (5.23) est une conséquence directe de la proposition 5.6.4. Notons que $\lambda_N(P) > 0$ si et seulement si les vecteurs $\xi_1, \ldots, \xi_N$ sont linéairement indépendants.

Une autre formulation de la formule (5.23) est instructive. Un calcul simple montre que

$${}^tTT = \begin{bmatrix} \xi_1 \cdot \xi_1 & \xi_1 \cdot \xi_2 & \cdots & \xi_1 \cdot \xi_N \\ \vdots & \vdots & \ddots & \vdots \\ \xi_N \cdot \xi_1 & \xi_N \cdot \xi_2 & & \xi_N \cdot \xi_N \end{bmatrix}.$$

Le déterminant de la matrice ${}^tTT$ s'appelle le ***déterminant de Gram*** des vecteurs $\xi_1, \ldots, \xi_N$; notons celui-ci par $\operatorname{gram}(\xi_1, \ldots, \xi_N)$. Alors, la formule (5.23) peut s'écrire comme suit :

$$\lambda_N(P) = \{\operatorname{gram}(\xi_1, \ldots, \xi_N)\}^{1/2}. \tag{5.24}$$

En effet, $\det({}^tTT) = (\det T)^2 = \operatorname{gram}(\xi_1, \ldots, \xi_N)$, ce qui donne (5.24) grâce à (5.23).

Notons de nouveau la conséquence suivante de (5.24) : *les vecteurs $\xi_1, \ldots, \xi_N$ dans $\mathbb{R}^N$ sont linéairement indépendants si et seulement si* $\operatorname{gram}(\xi_1, \ldots, \xi_N) > 0$ *; dans le cas de la dépendance linéaire des $\xi_1, \ldots, \xi_N$,* $\operatorname{gram}(\xi_1, \ldots, \xi_N) = 0$.

Considérons maintenant $P[\xi_1, \ldots, \xi_d]$ avec $d < N$; c'est un sous-ensemble d'un sous-espace vectoriel H de $\mathbb{R}^N$ de dimension $\leqslant d$. La formule (5.24) nous suggère de faire la définition suivante : posons

$$\sigma_d(P[\xi_1, \ldots, \xi_d]) = \{\operatorname{gram}(\xi_1, \ldots, \xi_d)\}^{1/2} \tag{5.25}$$

où $\operatorname{gram}(\xi_1, \ldots, \xi_d) = \det[\xi_i \cdot \xi_j]$, $1 \leqslant i, j \leqslant d$. Nous appellerons σ_d la ***mesure d-dimensionnelle*** de l'ensemble $P[\xi_1, \ldots, \xi_d] \subset \mathbb{R}^N$. On établit rapidement que les vecteurs $\xi_1, \ldots, \xi_d$ dans $\mathbb{R}^N$ sont linéairement indépendants si et seulement si $\operatorname{gram}(\xi_1, \ldots, \xi_d) > 0$; dans le cas contraire, $\operatorname{gram}(\xi_1, \ldots, \xi_d) = 0$.

Nous établissons maintenant une inégalité simple mais importante pour $\lambda_N(\varphi(A))$ où φ est une application générale de classe C^1. Cette inégalité est quelquefois appelée l'***inégalité de Sard*** ; nous utiliserons cette inégalité de manière essentielle pour obtenir la formule du changement des variables dans les intégrales multiples. La démonstration de cette inégalité donnée ici est basée sur celle de Walter [41, vol. II, p. 246].

5.6.6 Proposition

Soit U une partie ouverte de $\mathbb{R}^N$ et $\varphi : U \to \mathbb{R}^N$ une application de classe C^1. Alors

$$\lambda_N(\varphi(A)) \leqslant \int_A |\det \varphi'| \tag{5.26}$$

pour tout sous-ensemble A ouvert ou fermé dans U. Si $S = \{x \in U : \det \varphi'(x) = 0\}$ est l'ensemble des points singuliers de φ alors $\lambda_N(\varphi(S)) = 0$.

REMARQUES. Contrairement aux situations précédentes il n'est pas possible de démontrer que $\varphi(A)$ est borélien si A est borélien ; en fait, ceci n'est en général pas vrai ; néanmoins, on peut démontrer que $\varphi(A)$ est mesurable (cf. § 5.2.12) si A est mesurable (a fortiori, si A est borélien), et que (5.26) reste valable (sans hypothèses supplémentaires sur φ) pour tout A mesurable. Si A est compact, $\varphi(A)$ est compact (à cause de la continuité de φ), donc borélien. Comme tout A ouvert est une réunion dénombrable des cubes fermés, on en déduit que $\varphi(A)$ est une réunion dénombrable d'ensembles compacts, donc borélien, si A est ouvert. Si A est fermé dans U, A est une réunion dénombrable d'ensembles compacts d'où l'on déduit, de nouveau, que $\varphi(A)$ est borélien. Ce genre de raisonnement nous permet de conclure que $\varphi(A)$ est borélien pour beaucoup d'autres types d'ensembles boréliens A.

DÉMONSTRATION

Démontrons d'abord ce qui suit : soient $\xi \in U$, $T = \varphi'(\xi)$; soit $\{C_k\}$, $k = 1, 2, \ldots$, une suite de cubes fermés telle que ε_k, la longueur des côtés de C_k, tend vers 0 lorsque $k \to \infty$ et telle que $\xi \in C_k$, $k \geqslant 1$; alors

$$\liminf_{k\to\infty} \frac{\lambda_N(\varphi(C_k))}{\lambda_N(C_k)} \leqslant |\det T| . \tag{5.27}$$

La relation (5.27) si $\det T = 0$ est démontrée, en annexe, à la section 5.14 ; supposons donc que $\det T \neq 0$; supposons aussi, sans perte de généralité, que $\xi = 0$ et $\varphi(\xi) = 0$; alors

$$\varphi(x) = Tx + |x|\,\rho(x)$$

où $|\rho(x)| \to 0$ lorsque $|x| \to 0$, d'où

$$T^{-1}\varphi(x) = x + |x|\,T^{-1}\rho(x)$$

avec $|T^{-1}\rho(x)| \to 0$ lorsque $|x| \to 0$. Rappelons que C_k est un ensemble de la forme

$$C_k = \left\{x \in \mathbb{R}^N : |x - a_k| \leqslant \frac{\varepsilon_k}{2}\right\}$$

où $a_k \in \mathbb{R}^N$ avec $|a_k| \leqslant \varepsilon_k/2$ car $0 \in C_k$; donc, si $x \in C_k$, on a

$$|x| \leqslant |x - a_k| + |a_k| \leqslant \varepsilon_k\,.$$

Ainsi, si $x \in C_k$, on a

$$T^{-1}\varphi(x) = x + y$$

avec $|y| = |x|\,|T^{-1}\rho(x)| < \varepsilon_k\delta_k$ où $2\delta_k = \sup_{0\leqslant|x|\leqslant\varepsilon_k}|T^{-1}\rho(x)|$.

Notons que $\delta_k \to 0$ lorsque $k \to \infty$; on conclut que

$$T^{-1}\varphi(C_k) \subset C_k + \big(\,\big]-\varepsilon_k\delta_k\,,\,\varepsilon_k\delta_k\,\big[\,\big)^N = \widetilde{C}_k$$

où $\widetilde{C}_k$ est un cube ouvert dont les côtés sont de longueur $\varepsilon_k + 2\varepsilon_k\delta_k$; ainsi

$$\lambda_N\big(T^{-1}\varphi(C_k)\big) \leqslant \big(\varepsilon_k + 2\varepsilon_k\delta_k\big)^N\,.$$

Comme

$$\lambda_N\big(T^{-1}\varphi(C_k)\big) = |\det T^{-1}|\lambda_N\big(\varphi(C_k)\big) = \frac{1}{|\det T|}\,\lambda_N\big(\varphi(C_k)\big)$$

et

$$\lambda_N(C_k) = \varepsilon_k^N\,,$$

on obtient

$$\frac{\lambda_N\big(\varphi(C_k)\big)}{\lambda_N(C_k)} \leqslant |\det T|\big(1 + 2\delta_k\big)^N\,,$$

d'où l'on obtient (5.27) en laissant $k \to \infty$.

Passons maintenant à la preuve de (5.26), d'abord lorsque A est un cube fermé ; notre démonstration sera par l'absurde. Si (5.26) est en défaut pour un cube fermé A, de côtés de longueur a, on aura un $\varepsilon > 0$ tel que

$$\lambda_N\big(\varphi(A)\big) > \int_A |\det \varphi'| + \varepsilon\lambda_N(A)\,. \tag{5.28}$$

Subdivisons A en 2^N sous-cubes fermés $A_1, \ldots, A_{2^N}$, non empiétants, chacun ayant des côtés de longueur $a/2$; par l'additivité des deux côtés de (5.28), on conclut qu'il existe au moins un cube C_1 parmi les A_i pour lequel (5.28) est valable avec C_1 en place de A, car sinon

$$\varphi(A) = \varphi\left(\bigcup_{i=1}^{2^N} A_i\right) = \bigcup_{i=1}^{2^N} \varphi(A_i)$$

et on aurait, en contradiction avec 5.28,

$$\lambda_N(\varphi(A)) \leqslant \sum_{i=1}^{2^N} \lambda_N(\varphi(A_i))$$
$$\leqslant \sum_{i=1}^{2^N} \int_{A_i} |\det \varphi'| + \varepsilon \sum_{i=1}^{2^N} \lambda_N(A_i) = \int_A |\det \varphi'| + \varepsilon \lambda_N(A)$$

(les A_i étant non empiétants, $\lambda_N(A_i \cap A_j) = 0$, $i \neq j$). On peut continuer ce procédé de subdivision pour obtenir une suite de cubes fermés C_1, C_2, ... telle que $C_1 \supset C_2 \supset \ldots$ avec les côtés de C_k ayant la longueur $a/2^k$ et telle que (5.28) est vérifié avec C_k en place de A. Soit maintenant $\xi \in \bigcap_k C_k$; pour ce $\xi \in U$, on a une suite de cubes fermés associés C_k tels que, selon (5.28),

$$\liminf_{k\to\infty} \frac{\lambda_N(\varphi(C_k))}{\lambda_N(C_k)} \geqslant \varepsilon + |\det T| \tag{5.29}$$

car

$$\lim_{k\to\infty} \int_{C_k} |\det \varphi'| = |\det T|$$

étant donné la continuité de $\det \varphi'$ et le fait que $\varphi'(\xi) = T$. Mais (5.29) contredit (5.27). Ainsi, (5.26) est établie pour A un cube fermé. Comme tout ensemble ouvert est une réunion dénombrable de cubes fermés non empiétants (prop. 5.5.2), on voit que (5.26) reste valable aussi pour A ouvert (voir remarque ci-dessous). Si A est compact, $A \subset U$, il existe un ensemble ouvert B tel que $A \subset B \subset \overline{B} \subset U$ avec $\overline{B}$ compact ; si $M = \sup\{|\det \varphi'(x)| : x \in \overline{B}\}$ et V est un ensemble ouvert tel que $A \subset V \subset B$ et $\lambda_N(V \setminus A) < \varepsilon$ alors

$$\lambda_N(\varphi(A)) \leqslant \lambda_N(\varphi(V)) \leqslant \int_V |\det \varphi'| \leqslant \int_A |\det \varphi'| + M\varepsilon\,,$$

d'où, $\varepsilon > 0$ étant arbitraire, on déduit (5.26) pour A compact. Maintenant, si A est fermé dans U, on a que $A = F \cap U$, F étant fermé dans $\mathbb{R}^N$; comme U est une réunion dénombrable des cubes fermés C_1, C_2, ... (prop. 5.5.2), on a

$$A = F \cap U = \bigcup_{n=1}^{\infty} (F \cap C_n)$$

où chaque $F \cap C_n$ est compact. Posons $A_n = \bigcup_{j=1}^{n} (F \cap C_j)$; alors, chaque A_n est compact, $A_1 \subset A_2 \subset \ldots$, $\lim_n A_n = \bigcup_n A_n = A$ d'où

$$\varphi(A_1) \subset \varphi(A_2) \subset \ldots, \quad \varphi(A) = \varphi\left(\bigcup_{n=1}^{\infty} A_n\right) = \bigcup_{n=1}^{\infty} \varphi(A_n) = \lim_{n\to\infty} \varphi(A_n)\,,$$

et on déduit (5.26) pour A à partir de sa validité pour chaque A_n en laissant $n \to \infty$. Finalement, si

$$S = \{x \in U : \det \varphi'(x) = 0\}$$

alors S est un ensemble fermé dans U, et de (5.26), on a immédiatement que $\lambda_N(\varphi(S)) = 0$. ♦

REMARQUES. Dans la démonstration précédente, on a utilisé le fait que certains cubes fermés C_1, C_2, ... sont non empiétants ; ceci a comme conséquence que, pour $i \neq j$, $C_i \cap C_j$ est contenue dans un hyperplan du type $x_k =$ constante de sorte que $\lambda_N(C_i \cap C_j) = 0$ (cf. Ex. 7, § 5.4.5) ; on a alors

$$\lambda_N(C_1 \cup C_2 \cup \cdots \cup C_n) = \sum_{i=1}^{n} \lambda_N(C_i)$$

d'où

$$\lambda_N\left(\bigcup_{n=1}^{\infty} C_n\right) = \sum_{n=1}^{\infty} \lambda_N(C_n)$$

et de même

$$\int_{\bigcup_n C_n} p \, d\lambda_N = \sum_{n=1}^{\infty} \int_{C_n} p \, d\lambda_N , \quad p \geqslant 0 .$$

5.6.7 Ensembles négligeables dans $\mathbb{R}^N$

Nous pouvons maintenant généraliser les affirmations contenues dans les exercices 5, 6 et 7 du paragraphe 5.4.5 en démontrant que *toute sous-variété différentielle E dans $\mathbb{R}^N$ de classe C^1, de dimension $k < N$, est telle que* $\lambda_N(E) = 0$.

Démontrons d'abord que si $\psi : V \to \mathbb{R}^N$ est une application de classe C^1 définie dans une partie ouverte $V \subset \mathbb{R}^N$, $m < N$, alors $\lambda_N(\psi(V)) = 0$; en effet, posons

$$\varphi(x_1, \ldots, x_N) = \psi(x_1, \ldots, x_m), \quad (x_1, \ldots, x_N) \in V \times \mathbb{R}^{N-m} ;$$

l'application $\varphi : U \to \mathbb{R}^N$, $U = V \times \mathbb{R}^{N-m}$, est de classe C^1 avec $\det \varphi'(x) = 0$, $x \in U$, d'où $\lambda_N(\varphi(U)) = 0$; comme $\varphi(U) = \psi(V)$ on conclut que $\lambda_N(\psi(U)) = 0$. Comme E est la réunion dénombrable des ensembles de la forme $\psi(V)$ (voir ci-dessous), on conclut que $\lambda_N(E) = 0$.

Dans les considérations des mesures dans $\mathbb{R}^N$, on a souvent besoin de la propriété suivante de $\mathbb{R}^N$ (appelée ***propriété de Lindelöf***) : *de tout recouvrement ouvert d'un ensemble quelconque $A \subset \mathbb{R}^N$ on peut extraire un recouvrement dénombrable.*

Pour établir cette propriété, on considère les boules du type $B(\xi; 1/n)$, $n = 1, 2, \ldots$, $\xi = (\xi_1, \ldots, \xi_N)$ tel que chaque ξ_i est un nombre rationnel ; ces boules forment une famille dénombrable que nous rangeons dans la suite $B_1, B_2, \ldots$ Si $B(a; r)$ est une boule quelconque il existe une boule de type $B(\xi; 1/n)$ telle que $a \in B(\xi; 1/n) \subset B(a; r)$; il suffit de choisir n tel que $1/n < r/2$ et choisir ξ ensuite tel que $\|\xi - a\| < 1/n$. Soit $\{W_\alpha\}_{\alpha \in I}$ un recouvrement ouvert de A ; posons

$$J = \{n : \exists\, \alpha \in I \text{ tel que } B_n \subset W_\alpha\}\,.$$

Pour chaque $n \in J$, on a un α_n tel que $B_n \subset W_{\alpha_n}$; alors $A \subset \bigcup_n W_{\alpha_n}$; en effet, si $x \in A$, il existe un $\alpha \in I$ tel que $x \in W_\alpha$; donc, il existe un $r > 0$ tel que $B(x; r) \subset W_\alpha$ d'où l'existence d'un n avec $x \in B_n \subset B(x; r) \subset W_\alpha$; cet indice n est évidemment dans J ; alors W_{α_n} est tel que $x \in B_n \subset W_{\alpha_n}$, ce qui établit $A \subset \bigcup_n W_{\alpha_n}$.

Si E est une sous-variété différentielle dans $\mathbb{R}^N$ de classe C^1 et de dimension $k < N$, chaque point $x \in E$ possède un voisinage ouvert W_x tel que $W_x \cap E$ peut être paramétré par une application $\psi : V \to \mathbb{R}^N$ de classe C^1 avec $\psi(V) = W_x \cap E$. On a vu qu'alors $\lambda_N(W_x \cap E) = 0$; selon la propriété de Lindelöf, il existe une famille dénombrable parmi les W_x, soit $\{W_{x_n}\}$ telle que $E \subset \bigcup_n W_{x_n}$, d'où

$$\lambda_N(E) = \lambda_N\left(\bigcup_n E \cap W_{x_n}\right) \leqslant \sum_n \lambda_N(E \cap W_{x_n}) = 0\,.$$

Donnons maintenant un critère nécessaire et suffisant pour qu'une partie borélienne $E \subset \mathbb{R}^N$ soit un ensemble négligeable : *pour tout $\varepsilon > 0$, il doit exister une famille dénombrable de cubes fermés $\{C_n\}$, $n = 1, 2, \ldots$, telle que*

$$E \subset \bigcup_{n=1}^{\infty} C_n \quad \textit{et} \quad \sum_{n=1}^{\infty} \lambda_N(C_n) < \varepsilon\,. \tag{5.30}$$

En effet, si (5.30) est vérifiée, $\lambda_N(E) < \varepsilon$ pour tout $\varepsilon > 0$, donc $\lambda_N(E) = 0$; d'autre part, si $\lambda_N(E) = 0$, pour tout $\varepsilon > 0$, il existe un ensemble ouvert $V \supset E$ tel que $\lambda_N(V) < \varepsilon$ (la régularité extérieure de λ_N) ; mais $V = \bigcup_n C_n$ pour une famille dénombrable de cubes fermés $\{C_n\}$ non empiétants, ce qui donne

$$\varepsilon > \lambda_N(V) = \sum_{n=1}^{\infty} \lambda_N(C_n)\,,$$

c'est-à-dire (5.30).

En fait, on peut aller plus loin et ne plus prendre en compte l'hypothèse préliminaire que E est borélien ; la condition (5.30) entraîne que E est λ_N-mesurable et négligeable (mais non nécessairement borélien) ; nous ne donnons pas la preuve facile de ce fait ici.

Le critère (5.30) est souvent une méthode utile pour établir qu'un ensemble est négligeable ; on aurait pu se servir de cette méthode aussi pour démontrer la négligeabilité des sous-variétés différentielles de dimension $k < N$, de classe C^1, dans $\mathbb{R}^N$.

5.7 CHANGEMENT DE VARIABLES DANS LES INTÉGRALES

5.7.1 Introduction

Rappelons un énoncé exact élémentaire qui concerne une intégrale d'une fonction continue réelle f définie au moins sur un intervalle compact $[a, b]$ de $\mathbb{R}$ (avec $a < b$). On a

$$\int_a^b f(x)\,\mathrm{d}x = \int_c^d f\big(\varphi(u)\big)\varphi'(u)\,\mathrm{d}u\,; \tag{5.31}$$

si φ est continûment dérivable dans l'intervalle $J = [c, d]$ si $c < d$, ou $[d, c]$ si $d < c$, où $\varphi(c) = a$, $\varphi(d) = b$ et f est continue dans l'intervalle $K = \varphi(J)$.

Notons que, d'après le théorème des valeurs intermédiaires pour les fonctions continues, on a l'inclusion $[a, b] \subset K$; les hypothèses garantissent que $(f \circ \varphi)\varphi'$ donne une fonction continue dans J ; ainsi, les deux intégrales dans (5.31) sont bien définies. La preuve de (5.31) se fait par l'observation que si F est une primitive de f dans K, alors $G = F \circ \varphi$ est une primitive de $(f \circ \varphi)\varphi'$ dans J d'où

$$\int_c^d f\big(\varphi(u)\big)\varphi'(u)\,\mathrm{d}u = G(d) - G(c) = F(b) - F(a) = \int_a^b f(x)\,\mathrm{d}x\,.$$

Si φ est monotone, on aura

$$\varphi(J) = K = [a, b]\,;$$

si $\varphi' \geqslant 0$ (φ croissante), on aura $c < d$; par contre, si $\varphi' \leqslant 0$ (φ décroissante), on aura $c > d$ et la formule (5.31) devient

$$\int_a^b f(x)\,\mathrm{d}x = -\int_d^c f\big(\varphi(u)\big)\varphi'(u)\,\mathrm{d}u = \int_d^c f\big(\varphi(u)\big)\big|\varphi'(u)\big|\,\mathrm{d}u$$

ce qui équivaut à

$$\int_K f(x)\,\mathrm{d}x = \int_J f\big(\varphi(u)\big)\big|\varphi'(u)\big|\,\mathrm{d}u \tag{5.32}$$

avec $\varphi(J) = K$. La formule (5.32) est donc équivalente à (5.31) sous la condition supplémentaire que φ soit monotone dans l'intervalle J. C'est sous la forme (5.32) que nous généralisons la formule au cas des intégrales multiples dans $\mathbb{R}^N$; la condition de monotonicité de φ est remplacée par celle plus forte d'injectivité, mais autrement la formule (5.32) subsiste dans $\mathbb{R}^N$ sans aucun changement

sauf celui de remplacer $|\varphi'(u)|$ par $|\det \varphi'(u)|$. La généralisation de (5.31) aux intégrales multiples dans $\mathbb{R}^N$ exige d'autres développements que nous ne ferons pas ici (cf. sect. 5.13).

En pratique, on applique (5.31) comme suit : on écrit

$$x = \varphi(u), \quad dx = \varphi'(u)\,du$$

et ensuite on détermine c et d tels que $\varphi(c) = a$, $\varphi(d) = b$. Comme φ n'est pas supposée être injective, il peut y avoir plusieurs choix pour c et d ; la formule (5.31) reste valable pour chacun des choix, *pourvu* que φ et f remplissent les conditions signalées. Même si ces conditions peuvent être affaiblies de différentes manières, les exemples simples montrent qu'une application aveugle du formalisme peut donner des résultats absurdes (cf. Ex. 1, 2, sect. 5.10).

Notons une particularité de l'énoncé validant (5.31) : aucune condition d'injectivité n'est imposée sur φ ; de plus, si $c > d$, on doit interpréter la formule selon la convention

$$\int_c^d = -\int_d^c .$$

Nous donnons maintenant une des formulations courantes et utiles de la ***formule de changement de variables dans les intégrales multiples***.

5.7.2 Proposition

Soit $\varphi : U \to \mathbb{R}^N$ une application injective de classe C^1 où U est une partie ouverte de $\mathbb{R}^N$. Si $f : V \to \overline{\mathbb{R}}_+$ est une fonction borélienne avec $V = \varphi(U)$ alors

$$\int_V f(x)\,dx = \int_U f\big(\varphi(u)\big)\,\big|\det \varphi'(u)\big|\,du\,. \tag{5.33}$$

La formule (5.33) reste valable pour une fonction borélienne $f : V \to \overline{\mathbb{R}}$ pourvu que l'une des deux intégrales dans (5.33) existe ; dans ce cas, l'autre intégrale existe aussi et les deux sont égales.

REMARQUE. Notons que selon les remarques après la proposition 5.6.5, $V = \varphi(U)$ est borélien ; en fait, l'injectivité de φ entraîne que V est ouvert (théorème de l'invariance du domaine de Brouwer, § 4.2.7) ce qui est facile à établir si $\det \varphi' \neq 0$ partout, une hypothèse que nous ne faisons pas ici. En effet, si, de plus, $\det \varphi' \neq 0$ partout, φ serait un difféomorphisme de classe C^1 ; selon le théorème d'inversion locale, φ serait alors une application ouverte d'où $V = \varphi(U)$ est ouvert.

DÉMONSTRATION

Nous considérons d'abord le cas où $\det \varphi' \neq 0$ partout dans U. Dans ce cas, $\psi = \varphi^{-1} : V \to U$ est aussi une application injective de classe C^1, $V = \varphi(U)$ étant alors automatiquement ouvert dans $\mathbb{R}^N$; de plus, $\varphi(A)$ et $\varphi^{-1}(B)$ sont boréliens si $A \subset U$, $B \subset V$ sont boréliens car φ, φ^{-1} sont toutes deux

continues. Nous démontrons d'abord que l'inégalité (5.26) de la proposition 5.6.6 reste valable pour toute partie borélienne $A \subset U$; en effet, si W est une partie ouverte quelconque telle que $A \subset W \subset U$, alors on a $\varphi(A) \subset \varphi(W)$, d'où

$$\lambda_N(\varphi(A)) \leqslant \lambda_N(\varphi(W)) \leqslant \int_W |\det \varphi'| ;$$

si $\mu(A) = \int_A p$, $p = |\det \varphi'|$, μ définit une mesure de Radon d'où

$$\mu(A) = \sup\{\mu(W) : W \supset A,\ W \text{ ouvert}\} ;$$

en utilisant ceci, on obtient immédiatement

$$\lambda_N(\varphi(A)) \leqslant \int_A |\det \varphi'| , \quad A \subset U ,\ A \text{ borélien} .$$

Si $\varphi(A) = B \subset V$, alors on en déduit que

$$\int_V 1_B = \lambda_N(B) \leqslant \int_{\varphi^{-1}(B)} |\det \varphi'| = \int_U (1_B \circ \varphi)|\det \varphi'|$$

pour B borélien contenu dans V, d'où on a immédiatement

$$\int_V f(x)\,dx \leqslant \int_U (f \circ \varphi)|\det \varphi'| \tag{5.34}$$

pour toute fonction borélienne simple $f \geqslant 0$; un passage à la limite par une suite croissante de fonctions boréliennes simples établit (5.34) pour toute fonction f borélienne, $f : V \to \overline{\mathbb{R}}_+$ (cf. § 5.2.10). En appliquant l'inégalité (5.34) à l'application $\psi = \varphi^{-1} : V \to U$ on a, pour $g : U \to \overline{\mathbb{R}}_+$ borélienne, que

$$\int_U g \leqslant \int_V (g \circ \psi)|\det \psi'| . \tag{5.35}$$

Rappelons que $\varphi \circ \psi =$ identité de sorte que

$$(\varphi' \circ \psi) \cdot \psi' = I , \quad (\det \varphi' \circ \psi) \cdot (\det \psi') = 1 . \tag{5.36}$$

Si $g = (f \circ \varphi)|\det \varphi'|$ dans (5.35), on obtient de (5.34) et de (5.35), en nous servant de (5.36), que

$$\int_V f \leqslant \int_U g \leqslant \int_V (g \circ \psi)|\det \psi'| = \int_V f \tag{5.37}$$

car $(g \circ \psi)|\det \psi'| = (f \circ \varphi \circ \psi)|\det \varphi' \circ \psi|\,|\det \psi'| = f$. Ainsi, on doit avoir égalité partout dans (5.37) ce qui donne

$$\int_V f = \int_U (f \circ \varphi)|\det \varphi'| ,$$

c'est-à-dire (5.33) pour f borélienne, $f : V \to \overline{\mathbb{R}}_+$.

Passons maintenant au cas général (sans hypothèse que $\det\varphi' \neq 0$). Soit $S = \{x \in U : \det\varphi'(x) = 0\}$; on a vu dans la proposition 5.6.5 que $\lambda_N(\varphi(S)) = 0$ et que S est fermé dans U ; la restriction de φ à la partie ouverte $U \setminus S$ est une application de classe C^1 telle que $\det\varphi' \neq 0$ partout ; donc, de ce qui précède, on conclut que pour toute fonction borélienne $f : V \to \overline{\mathbb{R}}_+$ on aura d'abord

$$\int_{\varphi(U\setminus S)} f = \int_{U\setminus S} (f\circ\varphi)\left|\det\varphi'\right| ;$$

en tenant compte que $\varphi(U \setminus S) = \varphi(U) \setminus \varphi(S)$ (car φ est injective) et que $V = \varphi(U)$, $\lambda_N(\varphi(S)) = 0$, on a

$$\int_V f = \int_{\varphi(U\setminus S)} f + \int_{\varphi(S)} f = \int_{\varphi(U\setminus S)} f ;$$

comme $\det\varphi' = 0$ sur S, on a évidemment

$$\int_{U\setminus S} (f\circ\varphi)\left|\det\varphi'\right| = \int_U (f\circ\varphi)\left|\det\varphi'\right| ,$$

d'où

$$\int_V f = \int_U (f\circ\varphi)\left|\det\varphi'\right|$$

pour $f : V \to \mathbb{R}_+$ borélienne. Finalement, si $f : V \to \overline{\mathbb{R}}$, on écrit $f = f_+ - f_-$ avec $f_\pm : V \to \overline{\mathbb{R}}_+$ ce qui donne la seconde partie de la proposition 5.7.2. ♦

5.7.3 Remarques

Du point de vue de la manipulation, la formule (5.33) s'applique de la manière suivante : on pose $x = \varphi(u)$; alors $\mathrm{d}x = p(u)\,\mathrm{d}u$ où $p(u) = \left|\det\varphi'(u)\right|$. La difficulté principale dans l'application de cette formule réside dans la vérification que φ est injective dans U et ensuite dans la détermination précise de $V = \varphi(U)$. Soulignons encore le cas particulier de (5.33) où $f = 1_B$, $B \subset V$; alors

$$\lambda_N(B) = \int_{\varphi^{-1}(B)} \left|\det\varphi'(u)\right| \mathrm{d}u . \tag{5.38}$$

Aussi, si $B = \varphi(A)$, $A \subset U$, alors

$$\lambda_N(\varphi(A)) = \int_A \left|\det\varphi'(u)\right| \mathrm{d}u . \tag{5.39}$$

On a vu dans la démonstration de la proposition 5.7.2 que ce sont les formules équivalentes (5.38) et (5.39) qui donnent presque immédiatement la formule (5.33) du changement de variables. Les formules (5.38) et (5.39) permettent de calculer les volumes N-dimensionnels de B, respectivement de $\varphi(A)$, en calculant une intégrale appropriée ; lorsque φ est une application affine ($\varphi(u) = Tu + a$), (5.39) devient la proposition 5.6.4 qui dit que $\lambda_N(\varphi(A)) =$

$|\det T|\lambda_N(A)$; en effet, un calcul simple montre qu'alors $\varphi'(u) = T$, une fonction constante. Mais remarquons que la démonstration de la formule générale du changement de variables (5.33) est basée sur la formule particulière concernant les applications affines. La présente démonstration de la proposition 5.7.2 est basée sur l'inégalité simple (5.26) qui n'exige pas que la transformation φ soit injective ; si l'on admet la validité de (5.26) pour tout ensemble borélien A (avec φ seulement de classe C^1, non nécessairement injective), on a immédiatement

$$\int_V f \leqslant \int_U (f \circ \varphi)|\det \varphi'|$$

pour toute fonction borélienne $f : V \to \overline{\mathbb{R}}_+$; si, de plus, φ est injective, on peut renverser cette inégalité en passant à φ^{-1}, ce qui est l'idée de la démonstration donnée de la proposition 5.7.2 ; cette démarche, dans ce contexte, est due à J. T. Schwartz (« *The formula for change in variables in a multiple integral* », *American Mathematical Montlhy*, vol. 61 (1954), pp. 81-85).

Il est peut-être intéressant d'indiquer le raisonnement géométrique (heuristique) qui est sous-jacent à la formule de changement de variables (5.33). On a vu que cette formule équivaut essentiellement à la formule pour le volume N-dimensionnel de $\varphi(A)$ donnée dans (5.39). Si C est un « petit » cube fermé dans U centré à u de côté « du », $\varphi(C)$ peut être considéré comme un « presque-parallélépipède » dans V dont le volume est $|\det T|\,du$ avec $T = \varphi'(u)$; l'interprétation précise de ce point de vue est basée sur l'approximation de φ autour de u par une fonction affine : $\varphi(v) \sim \varphi(u) + T(v-u)$, $v \in \mathbb{R}^N$. La formulation détaillée de cette idée intuitive peut être faite de différentes manières, mais aucune d'elles n'est vraiment rapide ; la démonstration donnée complètement ci-dessus est une des démonstrations précises possibles. Il en existent plusieurs autres dans la littérature (cf. Dieudonné [13, tome 3, p. 155], Rudin [36, p. 206], Schwartz [38, tome IV, p. 67], Spivak [39, p. 67] et bien d'autres textes). La formule de changement de variables (5.33) est une formule fondamentale aussi bien du point de vue théorique que du point de vue pratique ; on peut dire que c'est une formule de base de la topologie différentielle. Dans la suite, on verra beaucoup d'applications de cette formule.

5.7.4 Formulation abstraite de la formule de changement de variables dans les intégrales

Soit (X, Σ, μ) un espace mesuré σ-fini ; si $(Y, \mathcal{F})$ est un espace borélien quelconque et $\varphi : X \to Y$ est une application mesurable, alors la formule

$$\nu(B) = \mu\big(\varphi^{-1}(B)\big)\,, \quad B \in \mathcal{F}\,,$$

définit une mesure positive dans l'espace Y, appelée ***mesure image*** de μ par φ. Même si μ est σ-finie et les espaces X, Y sont très simples, ν n'est pas nécessairement σ-finie ; considérer par exemple $X = Y = \mathbb{R}$ et $\varphi(x) \equiv 0$; si μ est la mesure de Lebesgue sur $\mathbb{R}$ définie dans la tribu borélienne Σ de X, ν est

la mesure de Dirac « du poids ∞ » concentrée à $y = 0$. Il est facile d'imposer une condition raisonnable sur φ pour que la mesure image $\mu\varphi^{-1} = \nu$ soit σ-finie ; mais, on peut donner une formule générale de changement de variables sans aucune hypothèse sur φ, sauf celle de la mesurabilité. En effet, on a

$$\int_Y f(y)\nu(\mathrm{d}y) = \int_X f\big(\varphi(x)\big)\mu(\mathrm{d}x) \tag{5.40}$$

pour toute fonction $f : Y \to \overline{\mathbb{R}}$ mesurable, ou bien $\geqslant 0$, ou bien ν-intégrable (ou bien telle que $f \circ \varphi$ est μ-intégrable). La démonstration de (5.40) est immédiate ; elle est évidemment vérifiée pour $f = 1_B$, $B \subset Y$, par la définition même de $\nu = \mu\varphi^{-1}$; ensuite on aura (5.40) pour $f \geqslant 0$, simple et mesurable, puis successivement pour $f \geqslant 0$ et mesurable et enfin pour f mesurable quelconque selon le schéma indiqué dans le paragraphe 5.2.10.

Indiquons la relation entre la formule générale (5.40), si facile à démontrer, et la formule très spéciale (5.33), dans $\mathbb{R}^N$, qui n'est pas démontrable aussi simplement. Supposons que X, Y sont des parties mesurables de $\mathbb{R}^N$, munies de la mesure de Lebesgue λ_N, et $\varphi : X \to Y$ est une application injective telle que $\varphi(A)$ est mesurable si $A \subset X$ est mesurable ou borélien ; posons

$$\mu(A) = \lambda_N\big(\varphi(A)\big)\,, \quad A \subset X\,.$$

On vérifie facilement que μ est une mesure positive dans la tribu borélienne Σ de X ; dans cette vérification, l'injectivité de φ joue un rôle essentiel. Evidemment, la mesure image de μ par φ dans Y (c'est-à-dire $\mu\varphi^{-1}$) est la mesure de Lebesgue λ_N ; selon la formule (5.40) on aura

$$\int_Y f(y)\,\mathrm{d}y = \int_X f\big(\varphi(x)\big)\mu(\mathrm{d}x)$$

pour toute fonction $f : Y \to \overline{\mathbb{R}}_+$ borélienne (par exemple, dans cette discussion illustrative, nous laissons de côté les questions de mesurabilité et celles de la possibilité évidente de généralisation au cas $f : Y \to \overline{\mathbb{R}}$). Si la mesure positive $\mu = \lambda_N \circ \varphi$ est telle que $\mu(A) = 0$ si $\lambda_N(A) = 0$ (continuité absolue de μ par rapport à λ_N), alors μ sera donnée par une fonction densité $p_\varphi \geqslant 0$ (cf. Ex. 27, § 5.2.15) par rapport à la mesure de Lebesgue dans X, de sorte que l'on aura

$$\int_Y f(y)\,\mathrm{d}y = \int_X f\big(\varphi(x)\big)p_\varphi(x)\,\mathrm{d}x\,. \tag{5.41}$$

On voit que (5.41) est exactement comme (5.33) sauf pour l'identification $p_\varphi = \big|\det \varphi'\big|$. On peut, en fait, démontrer maintenant directement que si φ est dérivable, alors p_φ est donné par $\big|\det \varphi'\big|$ (Rudin [37, p. 151]) [En effet, c'est la formule (5.27) avec lim inf remplacé par lim et $\leqslant$ remplacé par $=$]. Mais cette démonstration directe doit être appuyée par d'autres considérations théoriques.

Néanmoins, le raisonnement donné ci-dessus, amenant à (5.41), est général ; il s'applique dans les situations autres que celles où X, Y sont des parties

de $\mathbb{R}^N$; il se base sur l'unique condition d'injectivité de φ et de mesurabilité de $\varphi(A)$ pour A mesurable.

Même si φ n'est pas injective, (5.40) reste toujours valable mais la relation entre μ et ν reste alors un peu implicite. Néanmoins, cette formule aussi est fondamentale pour toutes les considérations de la théorie des probabilités et des statistiques (cf. sect. 5.13).

Remarquons, finalement, que la formule de changement de variables donnée dans la proposition 5.7.2 ne peut évidemment pas être valable pour toutes les transformations φ de classe C^1 qui ne sont pas nécessairement injectives ; il suffit d'avoir une application $\varphi : U \to V$ qui « recouvre » V deux (ou plusieurs) fois pour mettre (5.33) en défaut (cf. Ex. 3, sect. 5.10).

5.8 QUELQUES CAS IMPORTANTS DE CHANGEMENT DE VARIABLES DANS LES INTÉGRALES MULTIPLES

Il s'agit ici de spécialiser la formule (5.33) de la proposition 5.7.2 à un certain nombre d'exemples importants de φ. Dans les discussions suivantes les fonctions f à intégrer seront ou bien $\geqslant 0$ ou bien intégrables et boréliennes ; aussi les ensembles A seront implicitement boréliens ou mesurables.

5.8.1 Transformation affine

Si $\varphi(x) = Tx + a$, $T = [a_{ij}]$, $x \in \mathbb{R}^N$, on vérifie sans peine que

$$\varphi'(x) = T\,, \quad \det \varphi' = \det T\,,$$

d'où, en supposant que $\det T \neq 0$ (pour avoir l'injectivité de φ), on a

$$\int_{\varphi(A)} f(x)\,\mathrm{d}x = |\det T| \int_A f\big(\varphi(u)\big)\,\mathrm{d}u$$

avec $A \xrightarrow{\varphi} \varphi(A) \xrightarrow{f} \mathbb{R}$.

Notons deux cas spéciaux de cette formule : si T est une matrice orthogonale, on aura $|\det T| = 1$; si $Tx = (\lambda_1 x_1, \ldots, \lambda_N x_N)$, $\lambda_i \in \mathbb{R}^*$, $i = 1, \ldots, N$, $x = (x_1, \ldots, x_N)$, alors $|\det T| = |\lambda_1 \lambda_2 \cdots \lambda_N|$; en particulier si $\lambda_1 = \ldots = \lambda_N = \lambda$, on aura $|\det T| = |\lambda|^N$ (ici, T est une homothétie si $\lambda > 0$).

5.8.2 Transformation polaire dans $\mathbb{R}^2$

Ici $\varphi(r, \theta) = (r\cos\theta\,,\, r\sin\theta)$, $r > 0$, $-\pi < \theta < \pi$; du point de vue de la proposition 5.7.2, on doit prendre

$$U = \,]\,0\,,\,\infty\,[\,\times\,]\,-\pi\,,\,\pi\,[\,, \quad V = \varphi(U) = \mathbb{R}^2 \setminus \mathbb{R}_-$$

avec $\mathbb{R}_- = \{(x,0) : x \leqslant 0\}$. Un calcul simple montre que

$$\varphi'(r,\theta) = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}, \quad \det\varphi'(r,\theta) = r$$

d'où

$$\int_V f(x,y)\,dx\,dy = \int_U f(r\cos\theta\,,\, r\sin\theta)r\,dr\,d\theta\,,$$

ce que l'on écrit comme

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_0^{\infty}\int_{-\pi}^{\pi} f(r\cos\theta\,,\, r\sin\theta)r\,dr\,d\theta$$

en ajoutant les ensembles négligeables $\mathbb{R}_-$ et N aux intégrales sur V et sur U respectivement où N est tel que $[\,0\,,\,\infty[\,\times[\,-\pi\,,\,\pi\,] = U \cup N$. La bijectivité de $\varphi : U \to V$ était discuté dans la section 3.5. Au besoin, on peut changer $[\,-\pi\,,\,\pi\,]$ à $[\,0\,,\,2\pi\,]$ ou à un autre intervalle de longueur 2π.

5.8.3 Transformation cylindrique dans $\mathbb{R}^3$

Ici $\varphi(r,\theta,z) = (r\cos\theta,\, r\sin\theta,\, z)$, $\varphi : U \to V$ où

$$U = \,]\,0\,,\,\infty[\,\times\,]\,-\pi\,,\,\pi\,[\,\times\mathbb{R}, \quad V = \varphi(U) = \left(\mathbb{R}^2 \setminus \mathbb{R}_-\right)\times\mathbb{R}.$$

Un calcul simple donne

$$\det\varphi'(r,\theta,z) = r\,, \quad (r,\theta,z) \in U\,.$$

En ajoutant des ensembles négligeables appropriés à U et à V, on obtient

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y,z)\,dx\,dy\,dz = \int_{-\infty}^{\infty}\int_0^{\infty}\int_{-\pi}^{\pi} f(r\cos\theta,\, r\sin\theta,\, z)r\,dr\,d\theta\,dz\,.$$

La bijectivité de $\varphi : U \to V$ est une conséquence de celle de φ dans le paragraphe 5.8.2 ; la remarque concernant l'intervalle $[\,-\pi\,,\,\pi\,]$ est applicable ici aussi.

5.8.4 Transformation sphérique dans $\mathbb{R}^3$

Ici $\Phi(r,\theta,\varphi) = (x,y,z)$, $\Phi : U \to V$ avec

$$x = r\sin\theta\cos\varphi\,,\; y = r\sin\theta\sin\varphi\,,\; z = r\cos\theta\,,\; r > 0,\; 0 < \theta < \pi,\; -\pi < \varphi < \pi\,,$$

$$U = \,]\,0\,,\,\infty[\,\times\,]\,0\,,\,\pi\,[\,\times\,]\,-\pi\,,\,\pi\,[$$
$$V = \mathbb{R}^3 \setminus F\,, \quad F = \{(x,0,z) : x \leqslant 0,\, z \in \mathbb{R}\}\,.$$

Notons que F est un ensemble fermé dans $\mathbb{R}^3$ contenu dans le plan d'équation $y = 0$; donc $\lambda_3(F) = 0$. Un calcul direct de $\Phi(r,\theta,\varphi)$ nous donne

$$\det\Phi(r,\theta,\varphi) = r^2\sin\theta\,, \quad (r,\theta,\varphi) \in U\,.$$

En ajoutant les ensembles négligeables appropriés à U et à V, on obtient

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^{\infty}\int_0^{\pi}\int_{-\pi}^{\pi} f\big(\Phi(r,\theta,\varphi)\big) r^2 \sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\varphi\,.$$

La bijectivité de $\Phi : U \to V$ était affirmée dans le paragraphe 3.8.2 ; de nouveau, l'intervalle $[-\pi\,,\,\pi]$ pour φ peut être changé à $[0\,,\,2\pi]$ ou à un autre intervalle de longueur 2π.

5.8.5 Transformation sphérique dans $\mathbb{R}^N$

C'est une généralisation des transformations données dans les paragraphes 5.8.2 et 5.8.4. Ici $\varphi : U \to V$ où

$$\varphi(r,\theta_1,\ldots,\theta_{N-1}) = (x_1,\ldots,x_N)\,,$$
$$r>0,\ 0<\theta_i<\pi\ (1\leqslant i\leqslant N-2),\ -\pi<\theta_{N-1}<\pi\ ,$$

$$\begin{aligned} x_1 &= r\cos\theta_1\\ x_2 &= r\sin\theta_1\cos\theta_2,\ \ldots\\ x_{N-1} &= r\sin\theta_1\sin\theta_2\cdots\sin_{N-2}\cos\theta_{N-1}\\ x_N &= r\sin\theta_1\cdots\sin\theta_{N-2}\sin\theta_{N-1} \end{aligned}$$

$$U =]0,\infty[\ \times\ (]0,\pi[)^{N-2}\ \times\ ]-\pi\,,\,\pi[$$
$$V = \mathbb{R}^N \setminus F\,,\quad F = \big\{x=(x_1,\ldots,x_N) : x_{N-1}\leqslant 0\,,\ x_N = 0\big\}\,.$$

Notons que F est un ensemble fermé dans $\mathbb{R}^N$ contenu dans l'hyperplan d'équation $x_N = 0$; donc $\lambda_N(F) = 0$. Pour voir que φ est une bijection entre U et V nous pouvons raisonner comme suit : posons

$$r = \big(x_1^2+\cdots+x_N^2\big)^{1/2}\,,\quad s_i = \sin\theta_i\,,\quad c_i = \cos\theta_i\,,\quad 1\leqslant i\leqslant N-1\,;$$

soit $x=(x_1,\ldots,x_N)\in V$; alors $r>0$ et

$$x_{N-1}^2 + x_N^2 > 0\,,\quad -r<x_1<r\,;$$

il existe donc un unique $\theta_1\in\,]0\,,\,\pi[$ tel que $x_1 = rc_1$ d'où

$$x_2^2+\cdots+x_N^2 = r^2s_1^2\,,\quad -rs_1<x_2<rs_1\,;$$

on en déduit l'existence d'un unique $\theta_2\in\,]0\,,\,\pi[$ tel que $x_2 = rs_1c_2$, d'où

$$x_3^2+\cdots+x_N^2 = r^2s_1^2s_2^2\,,\quad -rs_1s_2<x_3<rs_1s_2\,.$$

En continuant ainsi, on détermine θ_i, $1\leqslant i\leqslant N-2$, avec

$$x_{N-1}^2+x_N^2 = r^2s_1^2s_2^2\cdots s_{N-2}^2 > 0\,,$$

d'où on obtient un unique $\theta_{N-1} \in]-\pi, \pi[$ tel que

$$x_{N-1} = r s_1 s_2 \dots s_{N-2} c_{N-1}, \quad x_N = r s_1 s_2 \cdots s_{N-2} s_{N-1}.$$

Ainsi, chaque $x \in V$ est l'image d'un unique élément $(r, \theta_1, \dots, \theta_{N-1}) \in U$ sous l'application φ ; de là, on déduit immédiatement la bijectivité de $\varphi : U \to V$. Il est clair que φ est de classe C^1 ; vérifions maintenant que

$$\det \varphi'(r, \theta_1, \dots, \theta_{N-1}) = r^{N-1} (\sin\theta_1)^{N-2} (\sin\theta_2)^{N-3} \cdots (\sin\theta_{N-2}), \tag{5.42}$$

ce qui est une généralisation évidente des formules du paragraphe 5.8.2 ($N = 2$) et du paragraphe 5.8.4 ($N = 3$).

Pour cette vérification, nous reprenons l'écriture $s_i = \sin\theta_i$, $c_i = \cos\theta_i$; un calcul direct montre que

$$\det \varphi'(r, \theta_1, \dots, \theta_{N-1}) = r^{N-1} D_N(s_1, \dots, s_{N-1}\,; c_1, \dots, c_{N-1})$$

où D_N est le déterminant de la matrice (N, N) suivante :

$$\begin{bmatrix} c_1 & -s_1 & 0 & \dots & 0 \\ s_1 c_2 & c_1 c_2 & -s_1 s_2 & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ s_1 \cdots s_{N-2} c_{N-1} & c_1 s_2 \cdots s_{N-2} c_{N-1} & \dots & & -s_1 \cdots s_{N-1} \\ s_1 \dots s_{N-1} & c_1 s_2 \cdots s_{N-1} & \dots & & s_1 \cdots s_{N-2} c_{N-1} \end{bmatrix}$$

Il suffit maintenant de vérifier que

$$D_N(s_1, \dots, s_{N-1}\,; c_1, \dots, c_{N-1}) = s_1^{N-2} s_2^{N-3} \cdots s_{N-2}.$$

En développant le déterminant D_N suivant sa première ligne, on obtient

$$D_N = c_1 \cdot c_1 s_1^{N-2} D_{N-1} + s_1 \cdot s_1^{N-1} D_{N-1} = s_1^{N-2} D_{N-1}$$

où $D_{N-1} = D_{N-1}(s_2, \dots, s_{N-1}\,; c_2, \dots, c_{N-1})$. On obtient maintenant la formule pour D_N par récurrence.

En ajoutant les ensembles négligeables appropriés à U et à V, on obtient

$$\int_{\mathbb{R}^N} f(x)\,\mathrm{d}x = \int_U f \circ \varphi(r, \theta_1, \dots, \theta_{N-1}) \cdot r^{N-1} (\sin\theta_1)^{N-2} \cdots (\sin\theta_{N-2})\,\mathrm{d}r\,\mathrm{d}\theta_1 \cdots \mathrm{d}\theta_{N-1}. \tag{5.43}$$

Notons une autre présentation de la transformation sphérique dans $\mathbb{R}^N$ qui est plus proche de cette donnée dans le paragraphe 5.8.4 pour $N = 3$:

$$\Phi(r, \theta_1, \dots, \theta_{N-2}, \varphi) = (x_1, \dots, x_N)$$

$$
\begin{aligned}
x_1 &= r \sin\theta_1 \cdots \sin\theta_{N-2} \cos\varphi \\
x_2 &= r \sin\theta_1 \cdots \sin\theta_{N-2} \sin\varphi \\
x_3 &= r \cos\theta_1 \sin\theta_2 \cdots \sin\theta_{N-2} \\
x_4 &= r \cos\theta_2 \sin\theta_3 \cdots \sin\theta_{N-2} \\
&\cdots \\
x_{N-1} &= r \cos\theta_{N-3} \sin\theta_{N-2} \\
x_N &= r \cos\theta_{N-2}
\end{aligned}
$$

avec $r > 0$, $-\pi < \varphi < \pi$, $0 < \theta_i < \pi$ $(i = 1, \ldots, N-2)$. Alors

$$
\left|\det \Phi'(r, \theta_1, \ldots, \theta_{N-2}, \varphi)\right| = r^{N-1} \sin\theta_1 (\sin\theta_2)^2 \cdots (\sin\theta_{N-2})^{N-2} \quad (5.42')
$$

ce qui se déduit directement de (5.42).

5.9 INTÉGRALES DES FONCTIONS VECTORIELLES

Pour différentes applications physiques, on a besoin de définir l'intégrale d'une fonction à valeur dans un espace vectoriel de dimension finie. Sans perte de généralité, on peut supposer que l'espace vectoriel concerné est $\mathbb{R}^M$. Soit (X, Σ, μ) un espace mesuré ; si $f : X \to \mathbb{R}^M$ est une fonction telle que $f = (f_1, \ldots, f_M)$ avec chaque composante $f_i : X \to \mathbb{R}$ une fonction μ-intégrable (on dira alors que f ***est μ-intégrable***), alors nous posons

$$
\int_X f \, d\mu = \left(\int_X f_1 \, d\mu, \ldots, \int_X f_M \, d\mu \right) .
$$

Autrement dit, nous intégrons f « composante par composante ». La plupart des propriétés de cette intégrale vectorielle sont immédiates à partir des propriétés des intégrales scalaires ; nous ne les énonçons pas ici explicitement. Notons seulement deux propriétés simples.

(i) Si $u : \mathbb{R}^M \to \mathbb{R}^P$ est une application linéaire alors, pour toute fonction μ-intégrable f, $u \circ f$ est une fonction μ-intégrable et

$$
u\left(\int_X f \, d\mu \right) = \int_X (u \circ f) \, d\mu . \quad (5.44)
$$

Nous laissons la preuve de (5.44) comme un exercice qui utilisera le fait que $u : \mathbb{R}^M \to \mathbb{R}^P$ est donné par

$$
u(y_1, \ldots, y_M) = \left(\sum_{j=1}^{M} u_{ij} y_j \right)_{1 \leqslant i \leqslant P} , \quad u_{ij} \in \mathbb{R}, \ \forall i, j .
$$

(ii) Si f est μ-intégrable alors

$$\left|\int_X f\,\mathrm{d}\mu\right| \leqslant \int_X |f|\,\mathrm{d}\mu \tag{5.45}$$

$$\left\|\int_X f\,\mathrm{d}\mu\right\| \leqslant \int_X \|f\|\,\mathrm{d}\mu\,. \tag{5.46}$$

La preuve de (5.45) est immédiate ; il suffit de se rappeler que

$$\left|(y_1, \ldots, y_M)\right| = \max_{1\leqslant i\leqslant M} |y_i|\,.$$

La preuve de (5.46) est moins immédiate, même si sa signification intuitive comme une généralisation de l'inégalité (du triangle) euclidienne dans $\mathbb{R}^M$ paraît évidente. Pour avoir une preuve mathématique de (5.46) il suffit de se rappeler que, si $y = (y_1, \ldots, y_M) \in \mathbb{R}^M$, alors il existe un vecteur $u = (u_1, \ldots, u_M)$ tel que

$$\|u\| = 1\,, \quad u \cdot y = \|y\|\,;$$

en effet,

$$u_i = \frac{y_i}{\|y\|}\,, \quad 1 \leqslant i \leqslant M\,,$$

donne un tel u si $y \neq 0$. Nous utilisons u ainsi défini pour avoir une application linéaire $\mathbb{R}^M \to \mathbb{R}$ donnée par

$$\xi \longmapsto u \cdot \xi\,, \quad \xi \in \mathbb{R}^M\,;$$

alors, selon (5.44), on aura, avec $y = \int_X f\,\mathrm{d}\mu$,

$$\left\|\int_X f\,\mathrm{d}\mu\right\| = \|y\| = u \cdot y = \int_X (u \cdot f)\,\mathrm{d}\mu \leqslant \int_X |u \cdot f|\,\mathrm{d}\mu \leqslant \int_X \|f\|\,\mathrm{d}\mu$$

car $|u \cdot \xi| \leqslant \|u\| \cdot \|\xi\| = \|\xi\|$; ceci établit (5.46).

Donnons maintenant quelques exemples concrets des intégrales vectorielles. Pour ces exemples, nous prenons $X = A \subset \mathbb{R}^N$ muni de la mesure de Lebesgue ; si $p : A \to \mathbb{R}_+$ est une fonction interprétée comme une densité, et $f : A \to \mathbb{R}^M$ est telle que pf est λ_N-intégrable, alors

$$\int_A pf\,\mathrm{d}\lambda_N = \left(\int_A pf_1\,\mathrm{d}\lambda_N, \ldots, \int_A pf_M\,\mathrm{d}\lambda_N\right)$$

où $f = (f_1, \ldots, f_M)$. Si

$$0 < m = \int_A p\,\mathrm{d}\lambda_N < \infty$$

($m =$ « la masse » de A), alors

$$\frac{1}{m}\int_A pf\,\mathrm{d}\lambda_N \tag{5.47}$$

est interprétée comme la ***moyenne*** de la quantité vectorielle f par rapport à la distribution de masse donnée par la densité p. En prenant $f(x) = x$, (5.47) donne le ***centre de gravité*** (ou ***barycentre***) de la distribution de masse donnée par la densité p ; un cas important de ceci est donné par $p = 1$ (densité uniforme) qui donne le ***barycentre géométrique*** de A, en supposant naturellement que $0 < \lambda_N(A) < \infty$.

Si $A \subset \mathbb{R}^3$, $f : A \to \mathbb{R}^3$, alors

$$\int_A \{x \times f(x)\}\, \mathrm{d}x \tag{5.48}$$

définit le ***moment***, autour de 0, du champ de force f dans A. De même, si $\omega \in \mathbb{R}^3$,

$$\int_A (\omega \times x)p(x)\, \mathrm{d}x \tag{5.49}$$

définit le ***moment angulaire*** du corps A (ayant une répartition de masse selon la densité p) tournant autour d'un axe de rotation ω passant par 0. Les intégrales scalaires de la forme

$$\int_A f(x) \cdot g(x)\, \mathrm{d}x$$

où f et g sont des fonctions vectorielles apparaissent aussi dans les problèmes de mécanique.

Mentionnons dans ce contexte deux types d'intégrales scalaires qui sont considérées souvent en physique et en mécanique. Si $A \subset \mathbb{R}^N$, $f : A \to \mathbb{R}$,

$$u(x) = \int_A \gamma(x - y)f(y)\, \mathrm{d}y$$

définit le ***potentiel newtonien*** au point x de la distribution de charge donnée par f où

$$\gamma(x) = \begin{cases} c_N \cdot \|x\|^{2-N} & \text{si } N \neq 2 \\ c \cdot \ln\|x\| & \text{si } N = 2 \end{cases}$$

(où c, c_N sont des constantes appropriées).

Si $A \subset \mathbb{R}^N$, $B \subset \mathbb{R}^N$, $p : A \to \mathbb{R}_+$, l'expression

$$I_B = \int_A d^2(x, B)p(x)\, \mathrm{d}x\,,$$

où $d(x, B)$ est la distance du point x de l'ensemble B (souvent B est une droite dans $\mathbb{R}^3$), est le ***moment d'inertie*** de la distribution de masse p par rapport à B. (Rappelons que $d(x, B) = \inf_{b \in B} \|x - b\|$.)

Pour illustrer la définition du barycentre, démontrons une ***règle de Guldin*** (pour le calcul d'un volume) : *soit E un ensemble dans le plan $(x, y) \in \mathbb{R}^2$, contenu dans le demi-espace défini par $y > 0$ et d'aire finie ; soit A l'ensemble de $\mathbb{R}^3$ engendré par la rotation de E autour de l'axe des x, alors*

$$\lambda_3(A) = \lambda_2(E)L$$

où L est la longueur du chemin parcouru par le barycentre géométrique de E lors de la rotation.

En effet,

$$A = \{(x, y, z) : y = t\cos\theta\,,\ z = t\sin\theta\,,\ 0 \leqslant \theta \leqslant 2\pi\,,\ (x, t) \in E\}\ ;$$

selon le théorème de Fubini,

$$\lambda_3(A) = \int_{-\infty}^{\infty} \lambda_2(A_x)\,\mathrm{d}x$$

où A_x est la x-section de A ; en utilisant la transformation polaire pour le calcul de $\lambda_2(A_x)$, on a

$$\lambda_2(A_x) = \int_{E_x} \left(\int_0^{2\pi} \mathrm{d}\theta\right) r\,\mathrm{d}r = 2\pi \int_{E_x} r\,\mathrm{d}r$$

d'où

$$\lambda_3(A) = 2\pi \int_{-\infty}^{\infty} \left(\int_{E_x} y\,\mathrm{d}y\right) \mathrm{d}x = \lambda_2(E) \cdot \frac{2\pi}{\lambda_2(E)} \iint_E y\,\mathrm{d}x\,\mathrm{d}y\,.$$

Or, le barycentre de E est donné par

$$\left(\frac{1}{\lambda_2(E)} \iint_E x\,\mathrm{d}x\,\mathrm{d}y\,,\ \frac{1}{\lambda_2(E)} \iint_E y\,\mathrm{d}x\,\mathrm{d}y\right)\ ;$$

compte tenu de l'expression pour $\lambda_3(A)$ obtenue ci-dessus, on a établi la règle de Guldin pour le volume de A. Dans les applications géométriques, on prendra comme axe de rotation une droite quelconque qui n'intersecte pas l'ensemble planaire E d'aire finie.

Le théorème de Fubini et la formule de changement de variables nous donnent tous les outils nécessaires pour le calcul d'un grand nombre d'intégrales dans $\mathbb{R}^N$. Les exercices de la section 5.10 sont souvent des exemples illustratifs pour montrer comment ces calculs sont effectués en pratique. Ils contiennent aussi quelques formules intégrales importantes pour nombreuses applications.

5.10 EXERCICES

1 L'application aveugle du formalisme de changement de variables

$$x = \frac{1}{u}, \quad \mathrm{d}x = -u^{-2}\,\mathrm{d}u$$

donne le résultat absurde

$$0 < \int_{-1}^{1} \frac{\mathrm{d}x}{x^2+1} \stackrel{?}{=} \int_{-1}^{1} \frac{-\mathrm{d}u}{u^2+1} < 0\,.$$

Noter que $\varphi(u) = 1/u$ pour $u \in [-1\,,\,1]$ n'a pas les propriétés postulées pour la validité de (5.31).

2 On vérifie que l'équation

$$\int_0^1 f(x)\,\mathrm{d}x = \int_0^{5\pi/2} f(\sin u)\cos u\,\mathrm{d}u$$

est juste si $f : [-1\,,\,1] \to \mathbb{R}$ est continue ; si f est continue et définie seulement sur $[0\,,\,1]$, le côté gauche est défini mais le côté droit n'a pas de sens ; si $f(x) = 1$ pour $x \in [0\,,\,1]$ et $f(x) = 1/(1-x^2)$ pour $x \in\,]-1\,,\,0[$ avec $f(-1) = 0$, le côté gauche vaut 1 et le côté droit donne une fonction non intégrable.

3 Posons $\varphi(u,v) = (u^2, v)$, $(u,v) \in \mathbb{R}^2$. Vérifier que $\left|\det \varphi'(u,v)\right| = 2|u|$. Si $U = \,]-1\,,\,1[\,\times\,]0\,,\,1[$, vérifier que $\varphi(U) = [0\,,\,1[\,\times\,]0\,,\,1[$ et

$$\lambda_2\big(\varphi(U)\big) = 1 < \int_U 2|u|\,\mathrm{d}u\,\mathrm{d}v = 2\,;$$

ceci est en accord avec l'inégalité de la proposition 5.6.6 ; noter que la proposition 5.7.2 ne s'applique pas car φ n'est pas injective sur U.

4 Vérifier que si $f : \mathbb{R}^N \to \mathbb{R}$ est intégrable alors

$$\int_{\mathbb{R}^N} f(tx)\,\mathrm{d}x = |t|^N \int_{\mathbb{R}^N} f(x)\,\mathrm{d}x\,, \quad t \neq 0,\, t \in \mathbb{R}$$

et

$$\int_{\mathbb{R}^N} f(\widetilde{\sigma}x)\,\mathrm{d}x = \int_{\mathbb{R}^N} f(x)\,\mathrm{d}x$$

où $\widetilde{\sigma}(x_1, \ldots, x_N) = \big(x_{\sigma(1)}, \ldots, x_{\sigma(N)}\big)$, σ étant une permutation de $\{1, \ldots, N\}$.

5 Soit $A = \big\{(x,y) \in \mathbb{R}^2 : (x/a)^2 + (y/b)^2 \leqslant 1\big\}$, $a > 0$, $b > 0$. Vérifier que $\lambda_2(A) = \pi ab$ en utilisant la transformation linéaire $T(u,v) = (au, bv)$.

6 Soit $A \subset \mathbb{R}^2$ un sous-ensemble de la forme

$$\theta_1 \leqslant \theta \leqslant \theta_2\,, \quad 0 \leqslant r \leqslant \rho(\theta)$$

où (r, θ) sont les coordonnées polaires dans $\mathbb{R}^2$. Vérifier que

$$\lambda_2(A) = \frac{1}{2}\int_{\theta_1}^{\theta_2} \rho^2(\theta)\,\mathrm{d}\theta\,.$$

En déduire que l'aire d'un secteur circulaire de rayon R et d'angle θ est égale à $(1/2)R^2\theta$.

7 Soit A la partie bornée de $\mathbb{R}^2$ délimitée par la ***lemniscate*** de Bernoulli d'équation cartésienne

$$\left(x^2 + y^2\right)^2 = 2a^2(x^2 - y^2)\,.$$

Voir que $\lambda_2(A) = 2a^2$ en se servant de la méthode de l'exercice 6 sur la partie de A continue dans $\{(x, y) : x \geqslant 0,\, y \geqslant 0\}$.

8 Vérifier que le volume d'un ellipsoïde dans $\mathbb{R}^3$ de demi-axes a, b, c est égal à $(4\pi/3)abc$ en utilisant la transformation linéaire $T(u, v, \omega) = (au, bv, c\omega)$ qui ramène le problème au calcul du volume d'une boule.

9 Soit

$$E = \left\{(x, y) \in \mathbb{R}^2 : a \leqslant x \leqslant b\,,\ 0 \leqslant g(x) \leqslant y \leqslant f(x)\right\}$$

et

$$A = \left\{(x, y, z) \in \mathbb{R}^3 : a \leqslant x \leqslant b\,,\ y = t\cos\theta\,,\ z = t\sin\theta\,,\ g(x) \leqslant t \leqslant f(x)\right\}.$$

Voir que A est obtenue par la rotation de E autour de l'axe des x. Vérifier que, pour $a \leqslant x \leqslant b$, $\lambda_2(A_x) = \pi\{f^2(x) - g^2(x)\}$ d'où

$$\lambda_3(A) = \pi\int_a^b \left\{f^2(x) - g^2(x)\right\}\mathrm{d}x\,.$$

10 Calculer les volumes suivants dans $\mathbb{R}^3$.

(i) $A = \{(x,y,z) : x^2+y^2 \leqslant z \leqslant x\}$; $\lambda_3(A) = \dfrac{\pi}{32}$.

(A est la partie bornée de $\mathbb{R}^3$ délimitée par les surfaces d'équations $z = x^2+y^2$, $z = x$; calculer $\lambda_3(A)$ par l'intégrale

$$\lambda_3(A) = \iint_E \{x - (x^2+y^2)\}\,dx\,dy$$

où E est le disque du plan (x,y) dont l'équation de la circonférence est $y^2 + x^2 - x = 0$; la transformation polaire donne alors

$$\lambda_3(A) = \iint_{\widetilde{E}} (r\cos\theta - r^2) r\,dr\,d\theta$$

avec $\widetilde{E} = \{(r,\theta) : -\pi/2 \leqslant 0 \leqslant \pi/2\,,\ 0 \leqslant r \leqslant \cos\theta\}$, d'où le résultat.)

(ii) $A = \{(x,y,z) : x^2+y^2+z^2 \leqslant a^2\,,\ z \geqslant a-h\}$ $\quad (a>0,\ 0 \leqslant h \leqslant a)$;

$\lambda_3(A) = \dfrac{1}{3}\pi h^2(3a-h)$.

(A est la partie de la boule fermée $\overline{B}(0;a)$ au-dessus du plan $z = a-h$; comme $\lambda_2(A_z) = \pi(a^2-z^2)$, on a

$$\lambda_3(A) = \int_{a-h}^{a} \pi(a^2-z^2)\,dz\,.)$$

(iii) $A = B \times [0\,,\,h]$, $\quad B \subset \mathbb{R}^2$; $\quad \lambda_3(A) = \lambda_2(B)h$.

(A est un cylindre de base B dans le plan (x,y) de hauteur h.)

(iv) $A = \left\{\left(\dfrac{zx}{h}\,,\,\dfrac{yz}{h}\,,\,z\right) : (x,y) \in B\,,\ 0 \leqslant z \leqslant h\right\}$, $\quad B \subset \mathbb{R}^2$;

$\lambda_3(A) = \dfrac{h}{3}\lambda_2(B)$.

(A est un cône de sommet $(0,0,0)$, de hauteur h, ayant la base B dans le plan $z = h$; on a $\lambda_2(A_z) = \lambda_2(B)(z/h)^2$ d'où $\lambda_3(A) = \int_0^h \lambda_2(A_z)\,dz$.)

(v) $A = \left\{(x,y,z) : x^2 + \left(\sqrt{y^2+z^2} - a\right)^2 \leqslant r^2\right\}$ $\quad (0<r<a)$;

$\lambda_3(A) = 2\pi^2 a r^2 = \dfrac{\pi^2}{4}(\rho_1+\rho_2)(\rho_2-\rho_1)^2$ avec $\rho_1 = a-r$, $\rho_2 = a+r$.

(A est le tore solide engendré par la rotation de $E = \{(x,y) : x^2+(y-a)^2 \leqslant r\}$ dans le plan (x,y) autour de l'axe des x (cf. sect. 5.9, règle de Guldin pour le calcul d'un volume) d'où $\lambda_3(A) = (\pi r^2)\cdot 2\pi a$; ρ_1 s'appelle le rayon intérieur et ρ_2 le rayon extérieur du tore.)

(vi) $A = \left\{(x,y,z) : x^2+y^2-ax \leqslant 0\,,\ x^2+y^2+z^2 \leqslant a^2\right\} \quad (a>0)$;

$$\lambda_3(A) = \frac{4}{3}a^3\left(\frac{\pi}{2}-\frac{2}{3}\right).$$

(A est l'intersection des parties renfermées par le cylindre $x^2+y^2=ax$ et la sphère $x^2+y^2+z^2=a^2$; on a

$$\lambda_3(A) = 2\iint_E \left(a^2-x^2-y^2\right)^{1/2} \mathrm{d}x\,\mathrm{d}y$$

avec $E = \left\{(x,y) : x^2+y^2-ax \leqslant 0\right\}$; faire ensuite une transformation polaire.)

11 (Fonctions beta et gamma, Ex. 18-19, § 5.2.15) On va établir que

$$B(s,t) = \frac{\Gamma(s)\Gamma(t)}{\Gamma(s+t)} \quad (s>0\,,\ t>0)$$

en raisonnant comme suit :

$$\Gamma(s) = 2\int_0^\infty x^{2s-1}\,\mathrm{e}^{-x^2}\,\mathrm{d}x$$

$$\Gamma(s)\Gamma(t) = 4\int_0^\infty\int_0^\infty x^{2s-1}y^{2t-1}\,\mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y$$

$$= 2\Gamma(s+t)\int_0^{\pi/2}\left(\sin\theta\right)^{2t-1}\left(\cos\theta\right)^{2s-1}\mathrm{d}\theta$$

en faisant une transformation polaire ; d'ici on obtient

$$\Gamma(s)\Gamma(t) = \Gamma(s+t)B(s,t)$$

en posant $y=\sin^2\theta$, etc. En déduire que, pour $\alpha>-1$, $\beta>-1$,

$$\int_0^{\pi/2}\left(\sin\theta\right)^\alpha\left(\cos\theta\right)^\beta\mathrm{d}\theta = \frac{1}{2}B\left(\frac{\alpha+1}{2}\,,\,\frac{\beta+1}{2}\right) = \frac{\Gamma\left(\frac{\alpha+1}{2}\right)\Gamma\left(\frac{\beta+1}{2}\right)}{2\Gamma\left(\frac{\alpha+\beta+2}{2}\right)}\,.$$

12 (Intégrale gaussienne dans $\mathbb{R}$) Vérifier que $I=\int_{-\infty}^{\infty}\mathrm{e}^{-x^2}\,\mathrm{d}x=\sqrt{\pi}$, en écrivant

$$I^2 = \iint_{\mathbb{R}^2}\mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y = \int_0^{2\pi}\mathrm{d}\theta\int_0^\infty r\,\mathrm{e}^{-r^2}\,\mathrm{d}r = \pi\,.$$

En déduire que $\Gamma(1/2)=\sqrt{\pi}$ et

$$\int_{-\infty}^{\infty}\mathrm{e}^{-x^2}\cdot x^{2\alpha}\,\mathrm{d}x = \Gamma\left(\alpha+\frac{1}{2}\right)\,,\quad \alpha>0\,.$$

Si

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-a)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}\ (a \in \mathbb{R},\ \sigma > 0),$$

vérifier alors que

$$\int_{-\infty}^{\infty} p(x)\,dx = 1, \quad \int_{-\infty}^{\infty} xp(x)\,dx = a, \quad \int_{-\infty}^{\infty} (x-a)^2 p(x)\,dx = \sigma^2,$$

$$\int_{-\infty}^{\infty} (x-a)^{2n} p(x)\,dx = \{(2n-1)(2n-3)\cdots 1\}\sigma^{2n}, \quad n = 1, 2, \ldots,$$

$$\int_{-\infty}^{\infty} (x-a)^{2n+1} p(x)\,dx = 0, \quad n = 0, 1, 2, \ldots$$

(Noter que

$$\Gamma\left(n+\frac{1}{2}\right) = \left(n-\frac{1}{2}\right)\Gamma\left(n-\frac{1}{2}\right) = \left(n-\frac{1}{2}\right)\cdots\frac{1}{2}\Gamma\left(\frac{1}{2}\right).)$$

REMARQUE. p s'appelle la ***densité gaussienne*** ou ***normale*** dans $\mathbb{R}$ de paramètres $a \in \mathbb{R}$ et $\sigma > 0$.

13 (Intégrales gaussiennes dans $\mathbb{R}^N$) Soit

$$p(x) = c\exp\left(-\frac{1}{2}\left(\sigma^{-1}(x-a)\right)\cdot(x-a)\right), \quad x \in \mathbb{R}^N$$

où $a = (a_1, \ldots, a_N) \in \mathbb{R}^N$, $\sigma = [\sigma_{ij}]$ est une matrice symétrique (N, N) de type strictement positif et $c = (2\pi)^{-N/2}(\det\sigma)^{-1/2}$; p s'appelle la ***densité gaussienne*** ou ***normale*** dans $\mathbb{R}^N$ de paramètres a et σ. Nous vérifions

(i) $\displaystyle\int_{\mathbb{R}^N} p(x)\,dx = 1$;

(ii) $\displaystyle\int_{\mathbb{R}^N} x_i p(x)\,dx = a_i,\ 1 \leqslant i \leqslant N$;

(iii) $\displaystyle\int_{\mathbb{R}^N} (x_i - a_i)(x_j - a_j)p(x)\,dx = \sigma_{ij},\ 1 \leqslant i, j \leqslant N.$

Pour (i), il suffit de considérer $a = 0$; soit T une matrice orthogonale telle que ${}^tT\sigma T = \operatorname{diag}[\alpha_1^2, \ldots, \alpha_N^2]$, $\alpha_1, \ldots, \alpha_N > 0$; on a $\det\sigma = (\alpha_1 \cdots \alpha_N)^2$; en posant $Tu = x$ on aura

$$\int_{\mathbb{R}^N} p(x)\,dx = c\int_{\mathbb{R}^N} \prod_{i=1}^{N} \exp\left(-\frac{u_i^2}{2\alpha_i^2}\right) du_1 \cdots du_N$$

$$= c\prod_{i=1}^{N} \int_{-\infty}^{\infty} \exp\left(-\frac{u_i^2}{2\alpha_i^2}\right) du_i = 1.$$

Pour (ii) et (iii) aussi, il suffit de considérer $a = 0$; avec $T = [t_{ij}]$ comme ci-dessus et $Tu = x$ on aura

$$\int_{\mathbb{R}^N} x_i p(x)\,\mathrm{d}x = c\int_{\mathbb{R}^N}\left(\sum_j t_{ij}u_j\right)\prod_{k=1}^{N}\exp\left(-\frac{u_k^2}{2\alpha_k^2}\right)\mathrm{d}u_1\cdots\mathrm{d}u_N = 0$$

$$\begin{aligned}\int_{\mathbb{R}^N} x_i x_j p(x)\,\mathrm{d}x &= c\int_{\mathbb{R}^N}\left(\sum_k t_{ik}u_k\right)\left(\sum_k t_{jk}u_k\right)\prod_{\ell=1}^{N}\exp\left(-\frac{u_\ell^2}{2\alpha_\ell^2}\right)\mathrm{d}u_1\cdots\mathrm{d}u_N\\ &= c\sum_{k,k'} t_{ik}t_{jk'}\int_{\mathbb{R}^N}(u_k u_{k'})\prod_{\ell=1}^{N}\exp\left(-\frac{u_\ell^2}{2\alpha_\ell^2}\right)\mathrm{d}u_1\cdots\mathrm{d}u_N\\ &= \sum_k t_{ik}t_{jk}\alpha_k^2 = \sigma_{ij}\,,\end{aligned}$$

car $\sigma = T\,\mathrm{diag}[\alpha_1^2, \ldots, \alpha_N^2]\,{}^tT$, après quelques petits calculs.

14 Soit

$$B_p^N(a;R) = \left\{x\in\mathbb{R}^N : \left(\sum_{i=1}^{N}|x_i - a_i|^p\right)^{1/p} < R\right\}$$

où $x = (x_1, \ldots, x_N)$, $a = (a_1, \ldots, a_N)$, $p > 0$, $R > 0$. Vérifions que

$$\lambda_N\{B_p^N(a;R)\} = R^N\cdot\frac{2^N\left\{\Gamma\left(\frac{1}{p}+1\right)\right\}^N}{\Gamma\left(\frac{N}{p}+1\right)}\,.$$

Les transformations $x \mapsto x + a$ et $x \mapsto Rx$ permettent de réduire le problème au cas où $a = 0$, $R = 1$. Posons $\alpha_p(N) = \lambda_N\{B_p^N(0;1)\}$; la x_N-section de $B_p^N(0;1)$ pour $x_N = t$ est $B_p^{N-1}\big(0;(1-|t|^p)^{1/p}\big)$ dont la mesure λ_{N-1} est $(1-|t|^p)^{(N-1)/p}\alpha_p(N-1)$; donc, pour $N \geqslant 2$,

$$\begin{aligned}\alpha_p(N) &= \alpha_p(N-1)\int_{-1}^{1}(1-|t|^p)^{(N-1)/p}\,\mathrm{d}t\\ &= \frac{2}{p}\,\alpha_p(N-1)B\big((N+p-1)/p,\,1/p\big)\\ &= 2\alpha_p(N-1)\frac{\Gamma\left(\frac{N-1}{p}+1\right)\Gamma\left(\frac{1}{p}+1\right)}{\Gamma\left(\frac{N}{p}+1\right)}\,.\end{aligned}$$

Puisque $\alpha_p(1) = 2$ (longueur de l'intervalle $]-1\,,\,1[$) on obtient, par récurrence,

$$\alpha_p(N) = \frac{2^N \left\{\Gamma\left(\frac{1}{p}+1\right)\right\}^N}{\Gamma\left(\frac{N}{p}+1\right)}\,.$$

On en déduit que

$$\alpha_1(N) = \frac{2^N}{N!}\,, \quad \alpha_2(N) = \frac{\pi^{N/2}}{\Gamma(N/2+1)}\,.$$

Notons que $\alpha_2(N)$ est le volume N-dimensionnel de la boule ouverte $B(0;1) \subset \mathbb{R}^N$ et que $\alpha_1(N)$ est le volume N-dimensionnel de l'ensemble ouvert polyédrique (octaèdre régulier si $N = 3$, carré si $N = 2$)

$$B_1(0;1) = \left\{x \in \mathbb{R}^N : \sum_{i=1}^N |x_i| < 1\right\}.$$

Comme conséquence on obtient que si S est le N-simplexe (triangle si $N = 2$, tétraèdre régulier si $N = 3$) donné par

$$S = \left\{x \in \mathbb{R}^N : x_i \geqslant 0\,,\, i = 1, \ldots, N\,,\, \sum_{i=1}^N x_i < 1\right\},$$

alors $\lambda_N(S) = 1/N!$ car $B_1(0;1)$ est la réunion de 2^N ensembles (deux à deux disjoints) de même volume N-dimensionnel que S.

Vérifier que

$$\lim_{p\to\infty} \lambda_N\{B_p^N(a;R)\} = (2R)^N\,, \quad \lim_{p\to 0} \lambda_N\{B_p^N(a;R)\} = 0$$

(pour la seconde, on utilise : $\bigcap_{p>0} B_p^N(0;1) = \{0\}$). Dessiner $B_p^2(0;1)$ pour $p = 1/2$, 1, 2 et pour $p \to \infty$ et interpréter les derniers résultats géométriquement.

15 Soit $f :]0\,,\,\infty[\to \overline{\mathbb{R}}_+$ une fonction borélienne ; vérifier que

$$\int_{\mathbb{R}^N} f(\|x\|)\,\mathrm{d}x = \alpha_N \int_0^\infty f(r) r^{N-1}\,\mathrm{d}r$$

où $\alpha_N = N\alpha_2(N) = 2\pi^{N/2}/\Gamma(N/2)$, $\alpha_2(N)$ étant le volume N-dimensionnel de la boule unité $B = B(0;1) \subset \mathbb{R}^N$ (cf. Ex. 14). En déduire que $\int_B \|x\|^{-p}\,\mathrm{d}x < \infty$ si et seulement si $p < N$ et que $\int_{B^c} \|x\|^{-p}\,\mathrm{d}x < \infty$ si et seulement si $p > N$ (utiliser le paragraphe 5.8.5 et l'exercice 11).

16 Soit $h : \mathbb{R}^3 \to \mathbb{R}$ une fonction borélienne telle que h est bornée dans chaque boule et $|h(x)| \leqslant C\|x\|^{-\alpha}$ pour $\|x\| > \rho$ où $\alpha > 2$, $C > 0$. Démontrer que

$$u(a) = \int_{\mathbb{R}^3} \frac{h(x)}{\|x - a\|}\, \mathrm{d}x$$

est bien défini pour tout $a \in \mathbb{R}^3$.

$$\Big(\int_{\mathbb{R}^3} \frac{|h(x)|}{\|x-a\|}\, \mathrm{d}x = \int_{\mathbb{R}^3} \frac{|h(y+a)|}{\|y\|}\, \mathrm{d}y$$

$$\leqslant M \int_B \frac{\mathrm{d}y}{\|y\|} + 2^\alpha C \int_{B^c} \frac{\mathrm{d}y}{\|y\|^{\alpha+1}} < \infty$$

où $B = B(0;R)$, $R > \max\big(2\|a\|, \|a\| + \rho\big)$, M une constante appropriée (cf. Ex. 15) ; u s'appelle le ***potentiel*** au point a d'une distribution de charge de densité h.)

17 Soient $\Omega = \{(x_1, \ldots, x_N) : x_1 + \cdots + x_N < 1,\ x_i > 0,\ 1 \leqslant i \leqslant N\}$ et $f :]0, 1[\to \overline{\mathbb{R}}_+$ une fonction borélienne. Démontrer que si les a_i sont > 0, alors

$$\int_\Omega f(x_1 + \cdots + x_N) x_1^{a_1 - 1} \cdots x_N^{a_N - 1}\, \mathrm{d}x_1 \cdots \mathrm{d}x_N = c \int_0^1 f(t) t^{a_1 + \cdots + \alpha_N - 1}\, \mathrm{d}t$$

où $c = \big(\Gamma(a_1) \cdots \Gamma(a_n)\big) / \big(\Gamma(a_1 + \cdots + a_N)\big)$; si un $a_i \leqslant 0$, l'intégrale de gauche est ∞ si f n'est pas négligeable. (Utiliser la transformation $\varphi(u_1, \ldots, u_N) = (x_1, \ldots, x_N)$ où $x_1 = u_1(1 - u_2)$, $x_2 = u_1 u_2 (1 - u_3)$, $\ldots$, $x_{N-1} = u_1 \cdots u_{N-1}(1 - u_N)$, $x_N = u_1 \cdots u_N$ en vérifiant que $\varphi : \big(]0, 1[\big)^N \to \Omega$ est bijective et que $\det \varphi'(u) = u_1^{N-1} u_2^{N-2} \cdots u_{N-1}$; pour cette dernière propriété, il est commode d'observer que $\varphi = \xi \circ \psi$,

$$\psi(u_1, \ldots, u_N) = (u_1, u_1 u_2, \ldots, u_1 u_2 \cdots u_N)$$

$$\xi(v_1, \ldots, v_n) = (v_1 - v_2, v_2 - v_3, \ldots, v_{N-1} - v_N, v_N)$$

de sorte que $\det \varphi'(u) = \det \xi'\big(\psi(u)\big) \det \psi'(u) = \det \psi'(u)$ car $\det \xi'(v) = 1$.)

18 Soient $f : [0, 1] \to \overline{\mathbb{R}}_+$ une fonction borélienne et $\Omega = \{x \in \mathbb{R}^N : Ax \cdot x \leqslant 1\}$ où A est une matrice symétrique (N, N) de type strictement positif. Démontrer que

$$\int_\Omega f(Ax \cdot x)\, \mathrm{d}x = \frac{\alpha_N}{\sqrt{\det A}} \int_0^1 f(r^2) r^{N-1}\, \mathrm{d}r$$

où α_N est comme dans l'exercice 15. (Soit T une matrice orthogonale telle que ${}^tTAT = \mathrm{diag}(\alpha_1^2, \ldots, \alpha_N^2)$; les transformations $Tu = x$ et $\alpha_1 u_1 = v_1$, $\ldots$, $\alpha_N u_N = v_N$ ramènent l'intégrale sur Ω à une intégrale sur la boule unité B de la forme $\int_B f\big(\|v\|^2\big)\, \mathrm{d}v \cdot \big(\det A\big)^{-1/2}$; utiliser l'exercice 15 ensuite.)

19 Démontrer que

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y = \pi\big(2\alpha-1\big)^2$$

où $f(x,y) = \operatorname{card}\big\{(m,n)\in\mathbb{Z}^2 : m^2+n^2 \leqslant x^2+y^2\big\}$ et $\alpha = \sum_{k\geqslant 0}\mathrm{e}^{-k^2}$. (Voir que l'intégrale

$$I = \sum_{m,n}\int_{R(m,n)} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y\,, \quad (m,n)\in\mathbb{Z}^2\,,$$

où $R(m,n) = \big\{(x,y) : m^2+n^2 \leqslant x^2+y^2\big\}$; une transformation polaire donne ensuite que $I = \sum_{m,n}\pi\,\mathrm{e}^{-(m^2+n^2)}$, d'où le résultat.)

20 Soit $E = \big\{(x,y)\in\mathbb{R}^2 : 0\leqslant x\leqslant 2\,,\ 0\leqslant y\leqslant x\big\}$. Démontrer que

$$\iint_E \big\{(x-y)^2+2(x+y)+1\big\}^{-1/2}\,\mathrm{d}x\,\mathrm{d}y = 2\ln 2 - \frac{1}{2}\,.$$

(La transformation $\varphi(u,v) = (x,y)$, $x = v(1+u)$, $y = u(1+v)$ ramène l'intégrale sur E à $\iint_D \mathrm{d}u\,\mathrm{d}v$ où $D = \big\{(u,v) : 0\leqslant u\leqslant 1\,,\ u\leqslant v\leqslant 2/(1+u)\big\}$ d'où le résultat. Pour compléter le raisonnement, il suffit de voir que si $U = \mathring{D}$, $V = \mathring{E}$ alors $\varphi : U\to V$ est une bijection, ce qui revient à la résolution unique des équations $x = v(1+u)$, $y = u(1+v)$ pour x, y, u, v strictement positifs et à la vérification que $\varphi(U) = V$; pour ce dernier, voir que $\varphi(\operatorname{Fr} U) = \operatorname{Fr} V$.)

21 Soient A, B deux parties mesurables de $\mathbb{R}^N$ telles que $0<\lambda_N(A)<\infty$, $0<\lambda_N(B)<\infty$; supposons que $\lambda_N(A\cap B) = 0$. Si $E = A\cup B$, démontrer que le barycentre géométrique de E est donné par $\big(\lambda_N(A)a+\lambda_N(B)b\big)/\big(\lambda_N(A)+\lambda_N(B)\big)$ où a est le barycentre géométrique de A et b est le barycentre géométrique de B.

22 Soient f, g deux fonctions réelles intégrables définies dans $\mathbb{R}^N$. Démontrer que

$$h(x) = \int_{\mathbb{R}^N} f(x-y)g(y)\,\mathrm{d}y$$

est défini p.p.(x) et h est une fonction intégrable. (Vérifier que

$$\iint_{\mathbb{R}^N\times\mathbb{R}^N} \big|f(x-y)\big|\,\big|g(y)\big|\,\mathrm{d}x\,\mathrm{d}y < \infty$$

en utilisant le théorème de Fubini).

REMARQUE. h s'appelle la ***convolution*** de f et g, et on la note par $h = f\star g$.

5.11 APPROXIMATION PAR DES FONCTIONS DE CLASSE C^∞

5.11.1 Généralités

Dans cette section nous démontrons qu'une fonction intégrable dans $\mathbb{R}^N$ peut être approchée par une fonction de classe C^∞. Nous établissons aussi le théorème de Weierstrass permettant l'approximation d'une fonction continue par une fonction polynomiale.

Soit $f : \mathbb{R}^N \to \mathbb{R}$; le ***support*** de f est l'ensemble $\overline{A}$ où $A = \{x : f(x) \neq 0\}$; on dit que f est une ***fonction à support compact*** si $\overline{A}$ est compact.

Dans la suite on aura besoin d'exemples de fonctions $f : \mathbb{R}^N \to \mathbb{R}$ de classe C^∞ à support compact. A cette fin, considérons la fonction suivante $\varphi : \mathbb{R} \to \mathbb{R}_+$ définie par

$$\varphi(x) = \begin{cases} \exp\left(-\dfrac{1}{x}\right) & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0. \end{cases}$$

Il est un exercice bien connu (remontant à Cauchy) que φ est de classe C^∞ ; pour le démontrer, il suffit d'établir par récurrence que

$$\varphi^{(n)}(x) = \mathrm{e}^{-1/x} p_n(x) \quad \text{si } x > 0$$

où p_n est un polynôme en x^{-1} ; alors, on aura $\varphi^{(n)}(0) = 0$ pour $n = 1, 2, \ldots$ d'où l'appartenance à la classe C^∞ de φ. Posons maintenant

$$\psi(x) = c \cdot \varphi\left(1 - \|x\|^2\right), \quad x \in \mathbb{R}^N,$$

où $c > 0$ est choisi de telle manière que

$$\int_{\mathbb{R}^N} \psi(x)\,\mathrm{d}x = 1 ;$$

notons que ψ est de classe C^∞, $\psi \geqslant 0$ et $\psi(x) = 0$ si $\|x\| \geqslant 1$ de sorte que ψ est de classe C^∞ dans $\mathbb{R}^N$ et à support compact. Si ψ_δ est définie par

$$\psi_\delta(x) = \delta^{-N} \psi\left(\frac{x}{\delta}\right), \quad x \in \mathbb{R}^N \ (\delta > 0)$$

on vérifie immédiatement que $\psi_\delta : \mathbb{R}^N \to \mathbb{R}_+$ est de classe C^∞ à support compact avec

$$\int_{\mathbb{R}^N} \psi_\delta(x)\,\mathrm{d}x = 1, \quad \psi_\delta(x) = 0 \text{ si } \|x\| \geqslant \delta .$$

Si $f; \mathbb{R}^N \to \mathbb{R}$ est une fonction ***localement intégrable*** dans le sens que

$$\int_B |f(x)|\,\mathrm{d}x < \infty$$

pour toute boule B dans $\mathbb{R}^N$, alors on peut définir la fonction $f_\delta : \mathbb{R}^N \to \mathbb{R}$ par la formule

$$f_\delta(x) = \int_{\mathbb{R}^N} f(x-y)\psi_\delta(y)\,\mathrm{d}y\,, \quad x \in \mathbb{R}^N\ ;$$

en effet,

$$\int_{\mathbb{R}^N} \bigl|f(x-y)\psi_\delta(y)\bigr|\,\mathrm{d}y \leqslant M \int_{B(0;\delta)} \bigl|f(x-y)\bigr|\,\mathrm{d}y = M \int_{B(x;\delta)} \bigl|f(u)\bigr|\,\mathrm{d}u < \infty\,,$$

où $M = \sup_y \psi_\delta(y)$, montre que f_δ est bien définie partout dans $\mathbb{R}^N$. D'autre part, on voit, par une transformation évidente, que

$$f_\delta(x) = \int_{\mathbb{R}^N} \psi_\delta(x-y)f(y)\,\mathrm{d}y\,, \quad x \in \mathbb{R}^N\,,$$

ce qui montre que f_δ est de classe C^∞ ; en effet,

$$\mathrm{D}^\alpha f_\delta(x) = \int_{\mathbb{R}^N} \mathrm{D}^\alpha\psi_\delta(x-y)f(y)\,\mathrm{d}y\,, \quad x \in \mathbb{R}^N\,,$$

pour $\alpha = (i_1, \ldots, i_n)$ quelconque ($\mathrm{D}^\alpha = \mathrm{D}_{i_1,\ldots,i_n}$ est la dérivation par rapport à x) ce que nous pouvons justifier par la proposition 5.3.3. Si F est le support de f, la formule

$$f_\delta(x) = \int_F \psi_\delta(x-y)f(y)\,\mathrm{d}y\,, \quad x \in \mathbb{R}^N\,,$$

montre que le support de f_δ est contenu dans

$$F_\delta = \bigl\{x : d(x,F) \leqslant \delta\bigr\}$$

car si $x \in F_\delta^c$, on a $\|x-y\| > \delta$ pour tout $y \in F$ de sorte que $\psi_\delta(x-y) = 0$, $y \in F$ d'où $f_\delta(x) = 0$ si $x \in F_\delta^c$.

Le procédé esquissé ci-dessus s'appelle la ***régularisation*** de f par *convolution* avec ψ_δ ; elle donne *une fonction f_δ de classe C^∞ à support compact si f est une fonction intégrable à support compact.* On va établir dans la suite que $\int |f_\delta - f| \to 0$ lorsque $\delta \to 0$ dès que f est intégrable. Pour cela, nous aurons besoin de deux lemmes ; aussi, il est commode d'introduire la notation suivante pour les fonctions $f : \mathbb{R}^N \to \mathbb{R}$ intégrables

$$\|f\|_1 = \int_{\mathbb{R}^N} \bigl|f(x)\bigr|\,\mathrm{d}x\,.$$

Alors

$$\|tf\|_1 = |t|\|f\|_1\,, \quad t \in \mathbb{R},$$

et

$$\|f+g\|_1 \leqslant \|f\|_1 + \|g\|_1$$

si f, g sont deux fonctions intégrables, définies dans $\mathbb{R}^N$ et à valeurs réelles.

5.11.2 Lemme

Soit $f : \mathbb{R}^N \to \mathbb{R}$ une fonction intégrable ; pour tout $\varepsilon > 0$, il existe une fonction $g = g_\varepsilon : \mathbb{R}^N \to \mathbb{R}$ continue à support compact telle que

$$\|f - g\|_1 = \int_{\mathbb{R}^N} |f(x) - g(x)| \, dx < \varepsilon .$$

DÉMONSTRATION

Il suffit d'établir le lemme pour $f \geqslant 0$, le cas général étant traité par la relation $f = f^+ - f^-$.

(i) Considérons d'abord le cas de $f = 1_A$, $\lambda(A) < \infty$ où $\lambda = \lambda_N$; si $\varepsilon' > 0$, il existe, selon le paragraphe 5.5.1, un ensemble compact K et un ensemble ouvert V tels que

$$K \subset A \subset V , \quad \lambda(V \setminus K) < \varepsilon' .$$

Selon le paragraphe 1.9.1, il existe $h : \mathbb{R}^N \to \mathbb{R}$ continue telle que

$$h(x) = \begin{cases} 1 & \text{si } x \in K \\ 0 & \text{si } x \in V^c, \end{cases} \qquad 0 \leqslant h \leqslant 1 ;$$

alors

$$\|1_A - h\|_1 = \int_{V \setminus K} |1_A - h| \leqslant 2\lambda(V \setminus K) < 2\varepsilon' .$$

Notons que si A est borné on pourra choisir V également borné et alors h sera continue à support compact contenu dans $\overline{V}$.

(ii) Soit f une fonction simple, $\geqslant 0$ et intégrable ; alors $f = \sum_{i=1}^n \alpha_i 1_{A_i}$ avec les $\alpha_i > 0$ et les $\lambda(A_i) < \infty$, $i = 1, \ldots, n$. Soit h_i une fonction continue telle que

$$0 \leqslant h_i \leqslant 1 , \quad \|1_{A_i} - h_i\|_1 < \varepsilon' , \quad i = 1, \ldots, n ;$$

alors, si $h = \sum_i \alpha_i h_i$, on aura

$$\|f - h\|_1 \leqslant \sum_i \alpha_i \|1_{A_i} - h_i\|_1 < \varepsilon' \cdot \sum_i \alpha_i .$$

Si f est à support compact contenu dans un ensemble ouvert borné V, les A_i seront contenus dans V de sorte que la fonction h sera continue à support compact contenu dans $\overline{V}$ selon la remarque de (i).

(iii) Soit $f : \mathbb{R}^N \to \mathbb{R}_+$ une fonction intégrable ; soit B une boule ouverte, $B = B(0; r)$, telle que (cf. Ex. 13, § 5.2.15) :

$$\int_{B^c} f(x) \, dx < \frac{\varepsilon}{3} .$$

Selon la définition même d'une intégrale, il existe une fonction simple, $a \geqslant 0$ et intégrable, telle que

$$0 \leqslant a \leqslant f\,1_B\,, \quad \|f\,1_B - a\|_1 < \frac{\varepsilon}{3}\,.$$

Evidemment, a est de la forme $\sum_{i=1}^n \alpha_i\,1_{A_i}$ avec $\alpha_i > 0$, $A_i \subset B$, $i = 1, \ldots, n$; de (ii), on sait qu'il existe une fonction continue g à support compact contenu dans $\overline{B}$ telle que

$$\|a - g\|_1 < \frac{\varepsilon}{3}\,;$$

il suffit de prendre ε' ci-dessus tel que $\varepsilon' \cdot \sum_i \alpha_i < \varepsilon/3$. Maintenant,

$$\begin{aligned}\|f - g\|_1 &= \int_{\overline{B}} |f - g| + \int_{\overline{B}^c} |f| \\ &\leqslant \int_{\overline{B}} |f - g| + \frac{\varepsilon}{3} \\ &\leqslant \int_{\overline{B}} |f - a| + \int_{\overline{B}} |a - g| + \frac{\varepsilon}{3} < \varepsilon\end{aligned}$$

puisque a et g sont nuls en dehors de $\overline{B}$ et $\lambda(\overline{B} \setminus B) = 0$. ♦

5.11.3 Lemme

Soit $f : \mathbb{R}^N \to \mathbb{R}$ une fonction intégrable ; pour $y \in \mathbb{R}^N$, posons $f_y(x) = f(x - y)$ et définissons

$$\alpha_f(y) = \|f_y - f\|_1 = \int_{\mathbb{R}^N} |f(x - y) - f(x)|\,\mathrm{d}x\,.$$

Alors $\alpha_f(y) \to 0$ lorsque $\|y\| \to 0$.

DÉMONSTRATION

Supposons d'abord que f est continue à support compact K ; alors

$$\alpha_f(y) = \int_{K \cup (K+y)} |f(x - y) - f(x)|\,\mathrm{d}x\,.$$

Puisque f est uniformément continue dans ce cas, pour tout $\varepsilon' > 0$, on peut choisir un $\delta > 0$ tel que

$$|f(x - y) - f(x)| < \varepsilon' \quad \text{si } \|y\| < \delta\,;$$

on aura alors, pour $\|y\| < \delta$,

$$\alpha_f(y) \leqslant \varepsilon' \lambda\big(K \cup (K + y)\big) \leqslant 2\varepsilon' \lambda(K)\,,$$

ce qui montre que $\alpha_f(y) \to 0$ lorsque $\|y\| \to 0$ dans le cas d'une fonction f continue à support compact. Pour f intégrable quelconque et $\varepsilon > 0$ donné, on peut prendre, selon le lemme 5.11.2, une fonction g continue à support compact tel que

$$\|f - g\|_1 < \varepsilon ;$$

alors

$$\begin{aligned}
\alpha_f(y) &= \int_{\mathbb{R}^N} |f(x-y) - g(x-y) + g(x-y) - g(x) + g(x) - f(x)| \, dx \\
&\leqslant \int_{\mathbb{R}^N} |f(x-y) - g(x-y)| \, dx + \int_{\mathbb{R}^N} |g(x-y) - g(x)| \, dx \\
&\quad + \int_{\mathbb{R}^N} |g(x) - f(x)| \, dx \\
&< 2\varepsilon + \alpha_g(y)
\end{aligned}$$

ce qui montre que

$$\limsup_{\|y\| \to 0} \alpha_f(y) \leqslant 2\varepsilon ;$$

$\varepsilon > 0$ étant arbitraire, le lemme est démontré. ♦

5.11.4 Proposition

Soit $f : \mathbb{R}^N \to \mathbb{R}$ une fonction intégrable ; si f_δ est la régularisation de f par convolution avec ψ_δ (§ 5.11.1), alors

$$\lim_{\delta \to 0} \|f - f_\delta\|_1 = \lim_{\delta \to 0} \int_{\mathbb{R}^N} |f(x) - f_\delta(x)| \, dx = 0 .$$

DÉMONSTRATION

On a

$$f_\delta(x) = \int_{\mathbb{R}^N} f(x-y)\psi_\delta(y) \, dy$$

où $\psi_\delta \geqslant 0$, $\psi_\delta(y) = 0$ si $\|y\| \geqslant \delta$, ψ_δ de classe C^∞ et $\int_{\mathbb{R}^N} \psi_\delta = 1$. Puisque f et ψ_δ sont intégrables, $f_\delta = f \star \psi_\delta$ est intégrable (Ex. 22, sect. 5.10) ; puisque $\int \psi_\delta = 1$, on a

$$f(x) = \int_{\mathbb{R}^N} f(x)\psi_\delta(y) \, dy ,$$

d'où

$$|f_\delta(x) - f(x)| \leqslant \int_{\mathbb{R}^N} |f(x-y) - f(x)| \psi_\delta(y) \, dy$$

et

$$\begin{aligned}
\|f - f_\delta\|_1 &\leqslant \int_{\mathbb{R}^N} \psi_\delta(y) \left(\int_{\mathbb{R}^N} |f(x-y) - f(x)| \, dx \right) dy \\
&= \int_{B_\delta} \psi_\delta(y) \, \alpha_f(y) \, dy
\end{aligned}$$

où $B_\delta = B(0;\delta)$ et α_f est défini dans le lemme 5.11.2 (car $\psi_\delta(y) = 0$ si $\|y\| \geqslant \delta$) ; pour $\varepsilon > 0$ donné, soit $\delta_0 > 0$ tel que $\alpha_f(y) < \varepsilon$ dès que $\|y\| < \delta_0$ (l'existence d'un tel δ_0 étant garantie par le lemme 5.11.2) ; on aura alors

$$\|f - f_\delta\|_1 \leqslant \varepsilon \int_{\mathbb{R}^N} \psi_\delta(y)\,\mathrm{d}y = \varepsilon$$

dès que $\delta \leqslant \delta_0$, ce qui établit la proposition. ♦

5.11.5 Proposition

Soit $f : \mathbb{R}^N \to \mathbb{R}$ une fonction intégrable ; pour tout $\varepsilon > 0$, il existe une fonction $g = g_\varepsilon : \mathbb{R}^N \to \mathbb{R}$ de classe C^∞ à support compact tel que

$$\|f - g\|_1 = \int_{\mathbb{R}^N} |f(x) - g(x)|\,\mathrm{d}x < \varepsilon\,.$$

DÉMONSTRATION

Soit B une boule ouverte, $B = B(0;r)$, telle que

$$\int_{B^c} |f| < \frac{\varepsilon}{2}\,.$$

Posons $h = f\,1_B$; alors h est une fonction intégrable à support compact contenu dans $\overline{B}$ telle que

$$\|f - h\|_1 = \int_{B^c} |f| < \frac{\varepsilon}{2}\,.$$

Selon la proposition 5.11.4, il existe un $\delta > 0$ tel que

$$\|h - h_\delta\|_1 < \frac{\varepsilon}{2}\,.$$

En posant $g = h_\delta$, on aura une fonction de classe C^∞ à support compact (cf. la discussion juste avant le lemme 5.11.2) telle que

$$\|f - g\|_1 \leqslant \|f - h\|_1 + \|h - g\|_1 < \varepsilon\,.$$ ♦

5.11.6 Proposition (*théorème de Weierstrass*)

Soit K une partie compacte de $\mathbb{R}^N$; si $f : K \to \mathbb{R}$ est une fonction continue et $\varepsilon > 0$, il existe une fonction polynomiale p de N variables $x_1, \ldots, x_N$ telle que

$$\sup_{x \in K} |f(x) - p(x)| \leqslant \varepsilon\,.$$

DÉMONSTRATION

Nous démontrons d'abord le théorème pour une fonction $f \in C(\mathbb{R}^N)$ telle que $f(x) = 0$ si $|x| \geqslant 1/2$ et pour $K \subset [-1/2\,,\,1/2]^N = P$.

Posons

$$p_n(x) = \alpha_n \prod_{i=1}^{N} (1 - x_i^2)^n, \quad x = (x_1, \ldots, x_N)$$

où $\alpha_n > 0$ est tel que

$$\int_P p_n(x)\,dx = 1.$$

Notons que p_n est un polynôme en $x_1, \ldots, x_N$ à coefficients réels tel que $p_n(x) \geqslant 0$ pour tout x dans P et

$$\lim_{n\to\infty} \int_{P_\delta} p_n(x)\,dx = 0 \tag{5.50}$$

où $P_\delta = \{x \in P : |x| \geqslant \delta\}$, $0 < \delta < 1$. En effet, si $0 < t < \delta < 1$, on a, pour tout $|x| \leqslant t$,

$$\prod_{i=1}^{N} (1 - x_i^2)^n \geqslant (1 - t^2)^{nN}$$

d'où

$$1 = \int_P p_n(x)\,dx \geqslant \alpha_n (1 - t^2)^{nN} \int_{|x|\leqslant t} dx = \alpha_n (2t)^N (1 - t^2)^{nN},$$

ce qui donne

$$\alpha_n \leqslant (2t)^{-N} (1 - t^2)^{-nN}.$$

Puisque

$$\prod_{i=1}^{N} (1 - x_i^2)^n \leqslant (1 - \delta^2)^{nN}$$

si $\delta \leqslant |x| \leqslant 1$, on a, pour tout $0 < t < \delta < 1$,

$$\int_{P_\delta} p_n(x)\,dx \leqslant \alpha_n (1 - \delta^2)^{nN} 2^N \leqslant \frac{2^N}{(2t)^N} \left(\frac{1 - \delta^2}{1 - t^2}\right)^{nN} \longrightarrow 0$$

lorsque $n \to \infty$, ce qui donne (5.50).

Posons

$$f_n(x) = \int_{\mathbb{R}^N} f(y) p_n(x - y)\,dy, \quad x \in \mathbb{R}^N.$$

Puisque $f(y) = 0$ si $|y| \geqslant 1/2$, l'existence de $f_n(x)$ est assurée ; de plus, p_n étant un polynôme, on voit que f_n est aussi un polynôme (ce qui se vérifie en écrivant $p_n(x-y)$ comme un polynôme en $y = (y_1, \ldots, y_N)$ dont les coefficients sont eux-mêmes des polynômes réels en $x_1, \ldots, x_N$). Si $|x| \leqslant 1/2$, on a

$$f_n(x) = \int_{\mathbb{R}^N} f(x - y) p_n(y)\,dy = \int_P f(x - y) p_n(y)\,dy$$

car $|y| > 1$ et $|x| \leqslant 1/2$ entraînent que $|x - y| > 1/2$ et $f(x - y) = 0$; donc, pour $|x| \leqslant 1/2$, on aura

$$f_n(x) - f(x) = \int_P \{f(x - y) - f(x)\} p_n(y)\, dy\,,$$

car $\int_P p_n = 1$, d'où

$$|f_n(x) - f(x)| \leqslant \int_P |f(x - y) - f(y)| p_n(y)\, dy$$

car $p_n \geqslant 0$ sur P. Ecrivons

$$P = P_\delta \cup (P \setminus P_\delta)\,.$$

Si δ est choisi tel que $0 < \delta < 1$ et

$$|f(x) - f(u)| < \frac{\varepsilon}{2}$$

dès que $|x - u| < \delta$ (ce qui est possible à cause de la continuité uniforme de f) on aura, pour $|x| \leqslant 1/2$,

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2} + 2M \int_{P_\delta} p_n(y)\, dy$$

où $M = \sup_{x \in \mathbb{R}^N} |f(x)|$; selon (5.50), on peut choisir n tel que

$$2M \int_{P_\delta} p_n(y)\, dy < \frac{\varepsilon}{2}\,;$$

en posant $f_n = p$, on aura alors

$$\sup_{x \in K} |f(x) - p(x)| \leqslant \sup_{|x| \leqslant 1/2} |f(x) - p(x)| < \varepsilon\,.$$

Passons maintenant au cas général d'une fonction $f \in C(K)$ où K est une partie compacte arbitraire de $\mathbb{R}^N$. Selon le théorème de prolongement de Tietze-Urysohn (§ 1.9.3), il existe une fonction $\tilde{f} \in C(\mathbb{R}^N)$ telle que

$$\tilde{f}(x) = f(x) \text{ si } x \in K\,, \quad \tilde{f}(x) = 0 \text{ si } |x| \geqslant T$$

où $T > 0$ est tel que $K \subset \{x : |x| < T\}$. En posant

$$g(x) = \tilde{f}(2Tx)\,, \quad x \in \mathbb{R}^N\,,$$

on aura une fonction $g \in C(\mathbb{R}^N)$ telle que $g(x) = 0$ si $|x| \geqslant 1/2$. Selon la première partie de cette démonstration, il existe un polynôme q tel que

$$\sup_{|x| \leqslant 1/2} |g(x) - q(x)| < \varepsilon\,.$$

Posons

$$p(x) = q\Big(\frac{x}{2T}\Big), \quad x \in \mathbb{R}^N ;$$

alors p est un polynôme tel que

$$\sup_{|x|\leqslant T} |\widetilde{f}(x) - p(x)| = \sup_{|x|\leqslant T} \Big|g\Big(\frac{x}{2T}\Big) - q\Big(\frac{x}{2T}\Big)\Big|$$
$$= \sup_{|y|\leqslant 1/2} |g(y) - q(y)| < \varepsilon ,$$

d'où

$$\sup_{x\in K} |f(x) - p(x)| \leqslant \sup_{|x|\leqslant T} |\widetilde{f}(x) - p(x)| < \varepsilon . \qquad \blacklozenge$$

5.11.7 Remarques

Les propositions 5.11.4 et 5.11.5 montrent qu'une fonction intégrable f quelconque peut être approchée « dans la norme $\mathcal{L}^1$ » par des fonctions de classe C^∞ et même par des fonctions C^∞ à support compact. En particulier, $\int_A f$ peut être approchée aussi précisément que l'on veut par $\int_A g$ où g est une fonction de classe C^∞ à support compact ; ainsi $\int_A f$ est ramenée, à peu de chose près, au calcul d'une intégrale « élémentaire » de la forme $\int_A g$. Malgré la grande généralité des fonctions intégrables, ces dernières permettent des approximations excellentes par des fonctions de classe C^∞ à support compact.

Le théorème de Weierstrass (prop. 5.11.6) est aussi de même nature ; malgré la grande généralité des fonctions continues, on peut les approcher uniformément sur chaque compact par les fonctions les plus élémentaires de l'analyse, c'est-à-dire par les fonctions polynomiales. La méthode de démonstration présentée ici est due à Landau ; un point de vue plus général nous amène au théorème connu comme ***théorème de Stone-Weierstrass*** qui est un des outils importants de l'analyse moderne avancée (cf. Dieudonné [13, tome I, p. 138], Schwartz « Analyse » [38, p. 340] entre autres).

5.12 QUELQUES COMPLÉMENTS

5.12.1 Inégalité de Jensen

Soit μ une mesure de probabilité dans un espace borélien (X, Σ) ; soient J un intervalle ouvert de $\mathbb{R}$ et $f : X \to J$ une fonction μ-intégrable ; si $\varphi : J \to \mathbb{R}$ est une fonction convexe telle que $\varphi \circ f$ est μ-intégrable, alors

$$\varphi\Big(\int_X f \,\mathrm{d}\mu\Big) \leqslant \int_X \varphi \circ f \,\mathrm{d}\mu .$$

DÉMONSTRATION

Posons $\alpha = \int_X f \, d\mu$; démontrons d'abord que $\alpha \in J$. Soit $J =]\alpha, b[$; démontrons que $\alpha > a$, la preuve que $\alpha < b$ étant similaire. Si $a = -\infty$, on n'a rien à vérifier ; supposons donc que a est un nombre fini. Selon l'hypothèse, $f(x) > a$ pour tout $x \in X$ ce qui donne

$$\alpha = \int_X f \, d\mu \geqslant a ;$$

si $\alpha = a$, on aurait

$$\int_X (f - a) \, d\mu = 0 ;$$

comme $f(x) - a > 0$ pour tout $x \in X$, ceci entraînerait que $\mu(X) = 0$ ce qui contredit $\mu(X) = 1$.

Rappelons (§ 1.9.4) que toute fonction convexe φ définie dans un intervalle ouvert est continue de sorte que $\varphi \circ f$ est une fonction mesurable ; du paragraphe 3.7.4, (3.34), on sait qu'il existe une constante réelle m telle que

$$\varphi(t) \geqslant \varphi(\alpha) + m(t - \alpha), \quad t \in J,$$

d'où on aura

$$\varphi\big(f(x)\big) \geqslant \varphi(\alpha) + m\big(f(x) - \alpha\big), \quad x \in X,$$

et

$$\int_X \varphi \circ f \, d\mu \geqslant \varphi(\alpha). \qquad \blacklozenge$$

REMARQUES. Un cas important de l'inégalité de Jensen est celui d'un ensemble borélien $X \subset \mathbb{R}^N$ avec $0 < \lambda_N(X) < \infty$; en prenant $\mu = \lambda_N / \lambda_N(X)$ on aura

$$\varphi\left(\frac{1}{\lambda_N(X)} \int_X f(x) \, dx\right) \leqslant \frac{1}{\lambda_N(X)} \int_X \varphi\big(f(x)\big) \, dx,$$

pourvu que φ soit convexe dans un intervalle ouvert approprié et que $\varphi \circ f$ soit intégrable dans X.

D'autres cas importants concernent des ensembles X dénombrables ou finis (comme $X = \mathbb{N}$ ou $\mathbb{Z}$ ou $\mathbb{N}^k$ ou $X = \{1, \ldots, n\}$) munis d'une fonction $p : X \to [0, 1[$ telle que $\sum_x p(x) = 1$; avec $X = \mathbb{N}$, par exemple, on aura

$$\varphi\left(\sum_{n=0}^{\infty} p_n x_n\right) \leqslant \sum_{n=0}^{\infty} p_n \varphi(x_n)$$

sous une hypothèse de convexité appropriée sur φ et la convergence de $\sum_n p_n |\varphi(x_n)|$ et de $\sum_n p_n |x_n|$.

Notons que la proposition 3.7.4 est aussi un cas particulier élémentaire où $X = \{1, \ldots, n\}$; cette inégalité était en fait démontrée précédemment dans

le paragraphe 1.9.4 à l'intérieur d'une discussion amenant vers la continuité des fonctions convexes dans un ensemble ouvert.

Comme mentionné à la fin du paragraphe 3.7.5, un grand nombre d'inégalités est une conséquence de l'inégalité de Jensen. Nous en donnons deux exemples maintenant.

5.12.2 Inégalités de Hölder et de Minkowski

Soit (X, Σ, μ) un espace mesuré quelconque. Si $0 < p < \infty$, on dit que $f \in \mathcal{L}^p(X)$ si $f : X \to \mathbb{R}$ est tel que f est mesurable et $|f|^p$ est μ-intégrable ; on pose alors

$$\|f\|_p = \left\{ \int_X |f|^p \, d\mu \right\}^{1/p} .$$

En fait, l'espace $\mathcal{L}^p(X)$ est considéré généralement seulement si $p \geqslant 1$. Démontrons d'abord l'***inégalité de Hölder*** : *si* $1 < p, q < \infty$, $1/p + 1/q = 1$ *et* $f \in \mathcal{L}^p(X)$, $g \in \mathcal{L}^q(X)$, *alors* $fg \in \mathcal{L}^1(X)$ *et*

$$\int_X |fg| \, d\mu \leqslant \|f\|_p \|g\|_q .$$

DÉMONSTRATION

De l'inégalité entre les moyennes arithmétiques et géométriques pondérées (Ex. 7, § 3.7.6), on a

$$a^\alpha b^\beta \leqslant \alpha a + \beta b$$

si $a \geqslant 0$, $b \geqslant 0$, $\alpha > 0$, $\beta > 0$, $\alpha + \beta = 1$ d'où en remplaçant a par a^p, b par b^q, α par $1/p$, β par $1/q$, on aura

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}$$

et

$$|f(x)g(x)| \leqslant \frac{1}{p} |f(x)|^p + \frac{1}{q} |g(x)|^q , \quad x \in X ,$$

ce qui montre que $fg \in \mathcal{L}^1(X)$; en supposant d'abord que $\|f\|_p = 1$ et $\|g\|_q = 1$, on obtient, en intégrant la dernière inégalité,

$$\int_X |fg| \, d\mu \leqslant 1 .$$

Si $\|f\|_p > 0$ et $\|g\|_q > 0$, on peut remplacer f par $f/\|f\|_p$ et g par $g/\|g\|_q$ dans l'inégalité précédente car

$$\left\| \frac{f}{\|f\|_p} \right\|_p = 1 , \quad \left\| \frac{g}{\|g\|_q} \right\|_q = 1$$

d'où résulte l'inégalité de Hölder signalée. ♦

Notons que le cas discret de l'inégalité de Hölder était démontré dans le paragraphe 3.7.5 par une méthode différente. Le cas $p = q = 2$ s'appelle aussi l'***inégalité de Cauchy-Schwarz***.

Nous déduisons de l'inégalité de Hölder l'***inégalité de Minkowski*** : *si* $1 \leqslant p < \infty$ *et* f, g *sont dans* $\mathcal{L}^p(X)$ *alors* $f + g \in \mathcal{L}^p(X)$ *et*

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p .$$

DÉMONSTRATION

Du fait que $x \mapsto x^p$ est convexe dans $[0, \infty[$ on déduit, grâce à l'inégalité de Jensen, que

$$\left(\frac{a+b}{2}\right)^p \leqslant \frac{a^p + b^p}{2}$$

si $a \geqslant 0$, $b \geqslant 0$, d'où

$$\left(|f(x)| + |g(x)|\right)^p \leqslant 2^{p-1} \left(|f(x)|^p + |g(x)|^p\right) , \quad x \in X ,$$

ce qui prouve que $(f + g) \in \mathcal{L}^p(X)$. Maintenant,

$$\begin{aligned} \left(\|f + g\|_p\right)^p &= \int_X |f + g|^p \, \mathrm{d}\mu \\ &\leqslant \int_X |f| \cdot |f + g|^{p-1} \, \mathrm{d}\mu + \int_X |g| \cdot |f + g|^{p-1} \, \mathrm{d}\mu \\ &\leqslant \left(\|f\|_p + \|g\|_p\right) \left(\int_X |f + g|^p \, \mathrm{d}\mu\right)^{1/q} \end{aligned}$$

en notant que si $q = p/(p-1)$ on aura $p^{-1} + q^{-1} = 1$ et en appliquant l'inégalité de Hölder aux couples $f \in \mathcal{L}^p(X)$, $|f+g|^{p-1} \in \mathcal{L}^q(X)$ et $g \in \mathcal{L}^p(X)$, $|f + g|^{p-1} \in \mathcal{L}^q(X)$; l'inégalité de Minkowski en résulte. ♦

REMARQUES. De l'inégalité de Minkowski on voit que $f \mapsto \|f\|_p$, $1 \leqslant p < \infty$, est une norme sur l'espace vectoriel $\mathcal{L}^p(X)$, sauf pour le fait que $\|f\|_p = 0$ n'entraîne pas que $f = 0$ mais seulement que $f = 0$ p.p.(μ) ; en identifiant deux fonctions qui sont égales p.p.(μ) on obtient une vraie norme $\|\cdot\|_p$ sur un espace vectoriel formé des classes d'équivalence des fonctions dans $\mathcal{L}^p(X)$. Nous ne nous en servirons pas ; seul le cas $p = 2$, $\mathcal{L}^2(X)$, appelé l'espace des ***fonctions de carré intégrable***, sera extrêmement important ; $\mathcal{L}^2(X)$ fournit un exemple concret d'un ***espace hilbertien***.

5.12.3 Formule d'Euler-Maclaurin, formule de Stirling

Nous établirons d'abord la formule suivante (***formule de Euler-Maclaurin***) : *si* $f : \mathbb{R}_+ \to \mathbb{R}$ *est de classe* C^1 *alors*

$$\sum_{k=0}^{n} f(k) = \int_0^n f(x)\,\mathrm{d}x + \frac{f(0)+f(n)}{2} + \int_0^n p(x)f'(x)\,\mathrm{d}x \tag{5.51}$$

où $p(x) = x - [x] - 1/2$, $x \in \mathbb{R}$ *et* $[x]$ = *la partie entière de* x.

($[x]$ est égal à l'unique entier m tel que $m \leqslant x < m+1$).

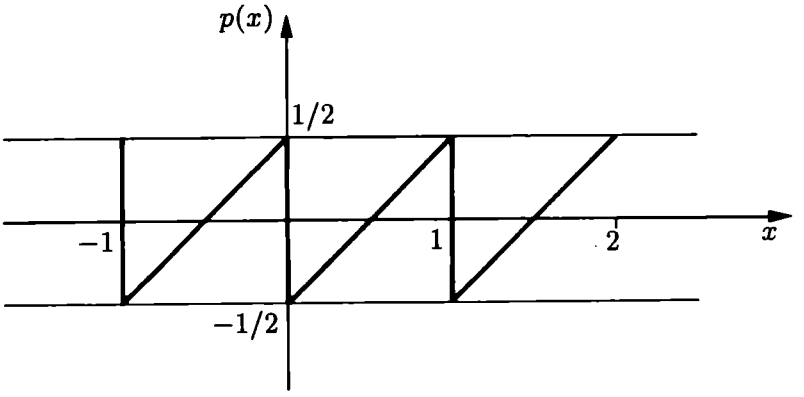


Fig. 5.3

DÉMONSTRATION

Notons que $p(x+1) = p(x)$, $x \in \mathbb{R}$; pour $k = 0, 1, \ldots$, on a

$$\begin{aligned}\int_k^{k+1} p(x)f'(x)\,\mathrm{d}x &= \int_k^{k+1} \left(x - k - \frac{1}{2}\right) f'(x)\,\mathrm{d}x \\ &= \left\{\left(x-k-\frac{1}{2}\right) f(x)\right\}_{x=k}^{x=k+1} - \int_k^{k+1} f(x)\,\mathrm{d}x \\ &= \frac{1}{2}\{f(k+1)+f(k)\} - \int_k^{k+1} f(x)\,\mathrm{d}x\ ;\end{aligned}$$

en sommant sur $k = 0, 1, \ldots, n-1$ on obtient (5.51). ♦

Nous allons déduire la ***formule de Stirling*** de (5.51) : *on a, pour* $n \to \infty$, $n \in \mathbb{N}^*$,

$$n! \sim \sqrt{2\pi}\,\mathrm{e}^{-n} n^{n+1/2} \tag{5.52}$$

où nous avons utilisé la notation suivante : si $\{a_n\}$, $\{b_n\}$ sont deux suites de nombres réels avec $b_n \neq 0$, $n = 1, 2, \ldots$, alors $a_n \sim b_n$ $(n \to \infty)$ équivaut à

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1\,.$$

On dit que $\{a_n\}$ est ***asymptotiquement équivalente*** à $\{b_n\}$.

DÉMONSTRATION

Posons $f(x) = \ln(1+x)$ dans (5.51) avec $n-1$ au lieu de n ; on aura alors

$$\sum_{k=1}^{n} \ln k = \int_1^n \ln x \, dx + \frac{1}{2} \ln n + \int_1^n \frac{p(x)}{x} \, dx$$

d'où, en utilisant $\int_1^n \ln x \, dx = n \ln n - (n-1)$,

$$\ln(n!) = \left(n + \frac{1}{2}\right) \ln n - (n-1) + \int_1^n \frac{p(x)}{x} \, dx .$$

Une intégration par parties donne

$$\int_1^n \frac{p(x)}{x} \, dx = \int_1^n \frac{P(x)}{x^2} \, dx$$

où

$$P(x) = \int_1^x p(t) \, dt , \quad x \geqslant 1 .$$

Noter que P est continue, $P(x+1) = P(x)$, $P(k) = 0$ si k est un entier et $P'(x) = p(x)$ pour x non entier ; en particulier, P est une fonction bornée. En posant

$$\alpha_n = 1 + \int_1^n \frac{p(x)}{x} \, dx = 1 + \int_1^n \frac{P(x)}{x^2} \, dx ,$$

on a

$$n! = e^{-n} n^{n+1/2} e^{\alpha_n} . \tag{5.53}$$

Puisque

$$\lim_{n \to \infty} \alpha_n = 1 + \int_1^{\infty} \frac{P(x)}{x^2} \, dx = \alpha ,$$

on en conclut que

$$n! \sim c e^{-n} n^{n+1/2}$$

où $c = e^{\alpha}$. Pour compléter la démonstration de la formule de Stirling, il faut vérifier que $c = \sqrt{2\pi}$. Nous le ferons en nous servant d'une astuce classique qui utilise la *formule de Wallis* :

$$\begin{aligned} \frac{\pi}{2} &= \lim_{n \to \infty} \left[\frac{2 \cdot 4 \cdots 2n}{1 \cdot 3 \cdots (2n-1)} \right]^2 \cdot \frac{1}{2n+1} \\ &= \lim_{n \to \infty} \left[\frac{2 \cdot 2 \cdot 4 \cdot 4 \cdots 2n \cdot 2n}{1 \cdot 3 \cdot 3 \cdot 5 \cdots (2n-1) \cdot (2n+1)} \right] . \end{aligned}$$

Nous établissons cette formule ci-après ; nous calculons d'abord

$$\lim_{n \to \infty} \binom{2n}{n} 2^{-2n} \sqrt{n}$$

de deux façons différentes, un fois à l'aide de (5.53) et une seconde fois en nous servant de la formule de Wallis. De (5.53), on a

$$\lim_{n\to\infty}\binom{2n}{n}2^{-2n}\sqrt{n}=\lim_{n\to\infty}\frac{2n!\sqrt{n}}{(n!)^2 2^{2n}}$$
$$=\lim_{n\to\infty}\sqrt{2}\exp(\alpha_{2n}-2\alpha_n)=\sqrt{2}\,\mathrm{e}^{-\alpha}$$

en utilisant le fait que $\alpha_n \to \alpha$ $(n \to \infty)$. D'autre part, puisque $(2n)! = \{1\cdot 3\cdots(2n-1)\}\cdot 2^n\cdot n!$, on a

$$\frac{2n!\sqrt{n}}{(n!)^2 2^{2n}}=\frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n}\sqrt{2n+1}\left(\frac{n}{2n+1}\right)^{1/2},$$

ce qui a comme limite, lorsque $n \to \infty$,

$$\sqrt{\frac{2}{\pi}}\cdot\frac{1}{\sqrt{2}}=\frac{1}{\sqrt{\pi}}$$

grâce à la formule de Wallis. Ainsi,

$$\sqrt{2}\,\mathrm{e}^{-\alpha}=\frac{1}{\sqrt{\pi}}$$

ce qui donne $c = \mathrm{e}^{\alpha} = \sqrt{2\pi}$.

Pour la démonstration de la formule de Wallis, on raisonne comme suit : posons

$$I_n=\int_0^{\pi/2}(\sin x)^n\,\mathrm{d}x,\quad n=1,2,\ldots$$

et notons que

$$I_{2n+1}<I_{2n}<I_{2n-1}\,.$$

Des formules données dans l'exercice 11 (sect. 5.10), on obtient

$$I_{2n}=\frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n}\frac{\pi}{2},\quad I_{2n+1}=\frac{2\cdot 4\cdots 2n}{3\cdot 5\cdots(2n+1)}\,.$$

Donc

$$\frac{2\cdot 4\cdots 2n}{3\cdot 5\cdots(2n+1)}<\frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n}\frac{\pi}{2}<\frac{2\cdot 4\cdots(2n-2)}{3\cdot 5\cdots(2n-1)}$$

d'où

$$\frac{2n}{2n+1}\frac{\pi}{2}<\left[\frac{2\cdot 4\cdots 2n}{1\cdot 3\cdots(2n-1)}\right]^2\cdot\frac{1}{2n+1}<\frac{\pi}{2},$$

ce qui donne la formule de Wallis. ♦

REMARQUES. La formule (5.51) est le cas le plus simple ($\nu = 0$) de la formule générale suivante due à Euler :

$$\sum_{k=0}^{n} f(k) = \int_0^n f(x)\,\mathrm{d}x + \frac{1}{2}\{f(0)+f(n)\} + \frac{B_2}{2!}\{f'(n)-f'(0)\} + \cdots$$
$$+ \frac{B_{2\nu}}{2\nu!}\{f^{(2\nu-1)}(n) - f^{(2\nu-1)}(0)\} + \int_0^n P_{2\nu+1}(x) f^{(2\nu+1)}(x)\,\mathrm{d}x$$

où f est de classe $C^{2\nu+1}$ dans $[0, \infty[$, les nombres B_n sont les ***nombres de Bernoulli*** définis par

$$\frac{x}{\mathrm{e}^x - 1} = \sum_{k=0}^{\infty} \frac{1}{k!} B_k\, x^k$$

($B_0 = 1$, $B_1 = -1$, $B_3 = B_5 = \ldots = 0$, $B_2 = 1/6$ $B_4 = -1/30$, ...) et les P_n sont les fonctions 1-périodiques ($P_n(x) = P_n(x+1)$, $x \in \mathbb{R}$, $n = 1, 2, \ldots$) et polynomiales dans $[0, 1[$ (***polynômes de Bernoulli***) définies par les relations

$$P_1(x) = x - \frac{1}{2}, \quad P'_{k+1}(x) = P_k(x), \quad k \geqslant 1,\ 0 < x < 1\ ;$$

$$P_{2k+1}(0) = 0\ (k \geqslant 0), \quad P_{2k}(0) = \frac{1}{2k!} B_{2k}\ (k \geqslant 1).$$

La fonction P_1 est la fonction p de (5.51) ; on vérifie, par exemple, que

$$P_2(x) = \frac{x^2}{2} - \frac{x}{2} + \frac{1}{12}, \quad P_3(x) = \frac{x^3}{6} - \frac{x^2}{4} + \frac{x}{12}\ ;$$

les nombres B_6, B_8, ... peuvent être calculés, par récurrence, mais ils deviennent très compliqués assez rapidement. Les nombres B_n et les polynômes P_n interviennent dans de nombreuses questions d'analyse (Knopp [23, p. 523]).

5.12.4 Exercices

1 En utilisant (5.51) avec $f(x) = (1+x)^{-1}$ démontrer que

$$\lim_{n\to\infty} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln n\right) = \gamma$$

où

$$\gamma = \frac{1}{2} - \int_1^{\infty} \frac{p(x)}{x^2}\,\mathrm{d}x$$

(γ s'appelle la ***constante d'Euler*** ; on sait que $\gamma = 0\cdot 5772\ldots$).

2 Si $\alpha > 1$, démontrer que

$$\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}} = \frac{1}{\alpha - 1} + \frac{1}{2} - \alpha \int_1^{\infty} \frac{p(x)}{x^{\alpha+1}}\,\mathrm{d}x$$

où p est la fonction de (5.51).

3 Si $0 < \alpha < 1$, démontrer que

$$\lim_{n\to\infty} \left\{ \sum_{k=1}^{n} \frac{1}{k^\alpha} - \frac{n^{1-\alpha}}{1-\alpha} \right\} = \frac{-1}{1-\alpha} + \frac{1}{2} - \alpha \int_1^\infty \frac{p(x)}{x^{\alpha+1}}\, \mathrm{d}x\,.$$

4 Soit (X, Σ, μ) un espace mesuré de mesure finie. Démontrer que si $0 < \alpha < \beta$ alors $f \in \mathcal{L}^\beta(X)$ entraîne que $f \in \mathcal{L}^\alpha(X)$ (utiliser le fait que si $|f(x)| > 1$ alors $|f(x)|^\alpha < |f(x)|^\beta$). Voir que cette affirmation est en défaut si $\mu(X) = \infty$ en considérant $X = [1, \infty[$ muni de la mesure de Lebesgue et $f(x) = 1/x$; $f \in \mathcal{L}^2(X)$ mais $f \notin \mathcal{L}^1(X)$.

5.13 COMMENTAIRES

La théorie de l'intégration dans un espace mesuré (X, Σ, μ) présentée dans la section 5.2 joue un rôle essentiel dans la théorie des probabilités ; dans cette théorie, $\mu(X) = 1$ et μ s'appelle une ***mesure*** (ou ***loi***) ***de probabilité*** ; une fonction mesurable $f : X \to \mathbb{R}^M$ s'appelle, dans cette théorie, une ***variable aléatoire*** (réelle si $M = 1$, vectorielle à valeurs dans $\mathbb{R}^M$ sinon) et son intégrale (si elle existe) s'appelle son ***espérance mathématique***. On écrit

$$\mathbb{E}(f) = \int_X f\, \mathrm{d}\mu\,.$$

La mesure de probabilité μf^{-1}, mesure image de μ par f (§ 5.7.4), s'appelle la ***loi*** de la variable aléatoire f ; la formule (5.40) s'écrit maintenant comme suit :

$$\mathbb{E}(f) = \int_{\mathbb{R}^M} y\nu(\mathrm{d}y)$$

où $\nu = \mu f^{-1}$ est la loi de la variable aléatoire f ; ν est une mesure de probabilité dans $\mathbb{R}^M$; si ν est donnée par une densité p on aura

$$\mathbb{E}(f) = \int_{\mathbb{R}^M} yp(y)\, \mathrm{d}y\ ;$$

si f prend seulement les valeurs $a_1, a_2, \ldots$ avec

$$p_n = \mu\{f^{-1}(a_n)\} = \mu\{x \in X : f(x) = a_n\}\,,$$

alors on aura

$$\mathbb{E}(f) = \sum_{n=1}^{\infty} a_n p_n\,.$$

Ainsi, la théorie des probabilités devient un chapitre particulier de la théorie de l'intégration, avec une terminologie spécifique déterminée par le développement historique du sujet. Le livre de Neveu [30] donne une présentation rigoureuse de la théorie mathématique du calcul des probabilités vu sous cet aspect, qui est dû à Kolmogorov (1933) dans sa forme définitive.

Il y a un développement différent de la théorie de l'intégration, popularisé par les écrits de Bourbaki, qui est aussi d'une très grande importance. Pour illustrer ce développement, considérons une partie $X \subset \mathbb{R}^N$, X étant soit compacte soit ouverte ; posons $C_0(X)$ pour l'espace des fonctions réelles, continues, à support compact et définies dans X ; évidemment, $C_0(X) = C(X)$ si X est compact. Soit

$$L : C_0(X) \to \mathbb{R}$$

une *fonctionnelle linéaire positive* ou une *forme linéaire positive* où « positive » signifie que $L(f) \geqslant 0$ si $f \geqslant 0$. Alors, un théorème important (appelé ***théorème de Riesz-Markov***) dit qu'*il existe une unique mesure de Radon positive μ définie dans la tribu borélienne Σ dans X telle que*

$$L(f) = \int_X f \, \mathrm{d}\mu \, ;$$

la mesure μ est nécessairement finie ($\mu(X) < \infty$) si X est compact ; si X est ouvert, μ est finie ou σ-finie. La mesure μ est une mesure de Radon dans le sens du paragraphe 5.5.1. Certain auteurs (comme Bourbaki) *définissent* une mesure de Radon positive dans X comme une forme linéaire positive $L : C_0(X) \to \mathbb{R}$. Le développement détaillé de cette théorie est un peu différent de celui présenté dans la section 5.2 ; néanmoins, les résultats finaux sont identiques, tout au moins dans le cas d'une partie $X \subset \mathbb{R}^N$, ouverte ou compacte. Ceci est assez clairement visible, à la vue du théorème de Riesz-Markov cité ci-dessus. De plus, on se rappellera que selon notre lemme 5.11.2, l'intégrale d'une fonction intégrable (par rapport à la mesure de Lebesgue) peut être approchée, aussi bien que l'on veut, par l'intégrale d'une fonction dans $C_0(X)$ ou même par l'intégrale d'une fonction dans $C_0(X)$ de classe C^∞ (prop. 5.11.4) ; cette possibilité d'approximation reste valable lorsque l'on considère l'intégration par rapport à une mesure de Radon μ quelconque dans $X \subset \mathbb{R}^N$. Voir, par exemple, Dieudonné [13, tome 2] ou Schwartz [38, tome 3] pour un développement du point de vue de Bourbaki sur la théorie de l'intégration.

La généralisation de la formule de changement de variables (sect. 5.7, (5.33)) aux cas où la transformation φ n'est pas injective ou n'est pas de classe C^1 pose des problèmes complexes ; une monographie difficile (Federer [18]) fait un tour complet du sujet (et de bien d'autres choses en relation avec notre livre) mais sa lecture est excessivement ardue ; une autre référence est Evans et Gariepy [17] (cf. les formules de « masses » et « co-masses » dans [17] et [18]).

5.14 APPENDICE

Soit $\varphi : U \to \mathbb{R}^N$ $(N = 1, 2, \ldots)$ une application de classe C^1, U étant une partie ouverte de $\mathbb{R}^N$; soit $T = \varphi'(\xi)$ où ξ est un point de U.

Si $\det T = 0$, *alors*

$$\lim_{k\to\infty} \frac{\lambda_N(\varphi(C_k))}{\lambda_N(C_k)} = 0$$

pour toute suite de cubes fermés $\{C_k\}_{k=1,2,\ldots}$ *telle que* $\xi \in C_k \subset U$, $k \geqslant 1$, *avec* $\varepsilon_k \to 0$, $k \to \infty$, *où* ε_k *est la longueur des côtés de* C_k. (λ_N = *mesure de Lebesgue dans* $\mathbb{R}^N$.)

DÉMONSTRATION

On peut supposer que $\xi = 0$ et $\varphi(\xi) = 0$; posons

$$C_k = \left\{ x \in \mathbb{R}^N : |x - a_k| \leqslant \frac{\varepsilon_k}{2} \right\}$$

où $|x|$ représente la norme maximum de $x = (x_1, \ldots, x_N)$: $|x| = \max_j |x_j|$; a_k est le centre du cube C_k. Puisque $\det T = 0$, il existe un sous-espace vectoriel H de dimension $(N-1)$ tel que $T(\mathbb{R}^N) \subset H$; soit $M > 0$ tel que $|Tx| \leqslant M|x|$ pour $x \in \mathbb{R}^N$. Puisque $0 \in C_k$, on voit que $|a_k| \leqslant \varepsilon_k/2$ et ensuite que

$$|x| \leqslant |x - a_k| + |a_k| \leqslant \varepsilon_k\,, \quad x \in C_k\,.$$

Ecrivons

$$\varphi(x) = Tx + |x|\,\rho(x)$$

où $|\rho(x)| \to 0$ lorsque $|x| \to 0$; si $A_k = T(C_k)$ on a $A_k \subset H$ et, pour $x \in A_k$, $|x| \leqslant M\varepsilon_k$ (car $x = Ty$, $y \in C_k$ et $|Ty| \leqslant M|y| \leqslant M\varepsilon_k$). Pour $\delta > 0$, on choisit k_0 tel que

$$|\rho(x)| < \delta\,, \quad x \in C_k\,,\ k \geqslant k_0\,,$$

ce qui est possible car, si $x \in C_k$, $|x| \leqslant \varepsilon_k \to 0$ $(k \to \infty)$; donc

$$|\varphi(x) - Tx| = |x|\,|\rho(x)| < \varepsilon_k\,\delta\,, \quad x \in C_k\,,\ k \geqslant k_0\,,$$

ce qui montre que

$$\varphi(C_k) \subset A_k(\varepsilon_k\delta)$$

où nous avons introtuit la notation suivante : si A est un ensemble dans $\mathbb{R}^N$, alors $A(t)$ est défini par

$$A(t) = \{ y \in \mathbb{R}^N : \exists\, x \in A \text{ t.q. } |x - y| < t \}\,.$$

Du lemme géométrique suivant on obtient que, pour tout $k \geqslant k_0$,

$$\lambda_N\big(\varphi(C_k)\big) \leqslant \lambda_N\big(A_k(\varepsilon_k\,\delta)\big) \leqslant N^{N/2}\varepsilon_k\,\delta\,\big(\varepsilon_k\,\delta + M\varepsilon_k\big)^{N-1}\,;$$

puisque $\lambda_N(C_k) = \varepsilon_k^N$ on a que

$$\frac{\lambda_N\big(\varphi(C_k)\big)}{\lambda_N(C_k)} \leqslant N^{N/2}\delta\big(\delta + M\big)^{N-1}, \quad k \geqslant k_0\,;$$

δ étant arbitraire on conclut que

$$\lim_{k\to\infty} \frac{\lambda_N\big(\varphi(C_k)\big)}{\lambda_N(C_k)} = 0 \qquad \blacklozenge$$

LEMME GÉOMÉTRIQUE

(Les notations ci-dessus sont maintenues.)

Soit

$$A \subset H \subset \mathbb{R}^N$$

où H est un sous-espace vectoriel de dimension $N-1$; si $\sup_{x\in A}|x| = d$ alors

$$\lambda_N\big(A(t)\big) \leqslant N^{N/2}t\big(t+d\big)^{N-1}, \quad t > 0\,.$$

DÉMONSTRATION

Soit $\{b_1, \ldots, b_N\}$ une base orthonormale de $\mathbb{R}^N$ telle que

$$H = \{x : x\cdot b_1 = 0\}, \quad b_2, \ldots, b_N \in H\,.$$

Pour $y \in \mathbb{R}^N$, écrivons $y = \sum_{i=1}^N \widetilde{y}_i b_i$, $\widetilde{y}_i = y\cdot b_i$; si $y \in A(t)$, il existe $x \in A$ tel que $|x-y| < t$ et

$$|\widetilde{y}_1| = |y\cdot b_1| = \big|(y-x)\cdot b_1\big| \leqslant \big\|y-x\big\| \leqslant \sqrt{N}|y-x| < t\sqrt{N}\,,$$
$$|\widetilde{y}_i| = |y\cdot b_i| \leqslant \big\|y\big\| \leqslant \sqrt{N}|y| \leqslant \sqrt{N}\big(|y-x| + |x|\big) < \sqrt{N}(t+d)\,.$$

Ainsi,

$$A(t) \subset P\big[t\sqrt{N}b_1, \sqrt{N}(t+d)b_2, \ldots, \sqrt{N}(t+d)b_N\big]$$

où $P[\xi_1, \ldots, \xi_N]$ est le parallélépipède de côtés $\xi_1, \ldots, \xi_N$ (sect. 5.6) ; on conclut que

$$\lambda_N\big(A(t)\big) \leqslant t\sqrt{N}\big(\sqrt{N}(t+d)\big)^{N-1} = N^{N/2}t\big(t+d\big)^{N-1}\,. \qquad \blacklozenge$$

CHAPITRE 6

Analyse vectorielle

L'objectif de ce chapitre est d'expliquer les trois théorèmes, appelés ici les théorèmes de Green, Gauss et Stokes respectivement, qui font le lien entre l'intégration et la dérivation dans $\mathbb{R}^2$ et $\mathbb{R}^3$. Le langage utilisé est celui des opérations vectorielles dans $\mathbb{R}^2$ et $\mathbb{R}^3$; mise à part la première section où l'intégrale curviligne est traitée dans $\mathbb{R}^N$, ce chapitre concerne l'analyse dans $\mathbb{R}^2$ et $\mathbb{R}^3$. Les démonstrations des trois théorèmes mentionnés ci-dessus, ne sont pas données dans ce chapitre ; elles seront reprises dans le chapitre suivant dans un cadre plus général utilisant le langage des formes différentielles. Le présent chapitre prépare le terrain pour une compréhension des développements du chapitre suivant et donne les considérations nécessaires pour l'étude des textes physiques et techniques qui sont toujours écrits, pour des raisons très valables, dans un langage vectoriel.

6.1 INTÉGRALES CURVILIGNES DANS $\mathbb{R}^N$

6.1.1 Définitions

Soit $\gamma : J \to \mathbb{R}^N$ une fonction continue définie dans un intervalle J de $\mathbb{R}$. Si $f : E \to \mathbb{R}$ est une fonction réelle définie dans $E \supset \gamma(J)$, on définit

$$\int_\gamma f \, \mathrm{d}s$$

par la formule

$$\int_\gamma f \, \mathrm{d}s = \int_J f\big(\gamma(t)\big) \big\|\gamma'(t)\big\| \, \mathrm{d}t \tag{6.1}$$

en supposant que $\gamma'(t)$ existe p.p. dans J et que l'intégrale sur J dans (6.1) existe au sens de la théorie de l'intégration de Lebesgue donnée dans le chapitre 5. De même, si $F : E \to \mathbb{R}^N$ est une fonction vectorielle définie dans $E \supset \gamma(J)$, on définit

$$\int_\gamma F \cdot \mathrm{d}\vec{s}$$

par la formule

$$\int_\gamma F \cdot \mathrm{d}\vec{s} = \int_J F\big(\gamma(t)\big) \cdot \gamma'(t) \, \mathrm{d}t \tag{6.2}$$

en supposant que $\gamma'(t)$ existe p.p. dans J et que l'intégrale sur J dans (6.2) existe au sens de Lebesgue.

L'intégrale

$$\int_\gamma f \, \mathrm{d}s$$

s'appelle l'***intégrale curviligne*** de la fonction scalaire (ou ***champ scalaire***) f le long de (la ***trajectoire***) γ.

L'intégrale

$$\int_\gamma F \cdot \mathrm{d}\vec{s}$$

s'appelle l'***intégrale curviligne*** du ***champ vectoriel*** F le long de γ ; une autre notation pour cette intégrale est la suivante :

$$\int_\gamma F \cdot \mathrm{d}\vec{s} = \int_\gamma \{F_1 \, \mathrm{d}x_1 + \cdots + F_N \, \mathrm{d}x_N\} \tag{6.3}$$

où $F = (F_1, \ldots, F_N)$; une justification immédiate pour cette notation est que si $\gamma = (\gamma_1, \ldots, \gamma_N)$, on aura, en posant

$$x_1 = \gamma_1 \,, \ \mathrm{d}x_1 = \gamma_1' \, \mathrm{d}t \,; \quad \ldots \,; \quad x_N = \gamma_N \,, \ \mathrm{d}x_N = \gamma_N' \, \mathrm{d}t \,,$$

l'égalité formelle

$$F_1 \, \mathrm{d}x_1 + \cdots + F_N \, \mathrm{d}x_N = F \cdot \gamma' \, \mathrm{d}t \,.$$

L'expression $\omega = F_1 \, \mathrm{d}x_1 + \cdots + F_N \, \mathrm{d}x_N$ s'appelle aussi une ***forme différentielle d'ordre*** 1 et l'on dit que $\int_\gamma F \cdot \mathrm{d}\vec{s}$ est l'intégrale de la forme ω le long de γ.

Dans la suite, l'existence des intégrales sur J dans (6.1) et (6.2) sera toujours évidente ; en effet, on aura (presque toujours) un intervalle compact J avec f, F continues et γ une fonction dérivable sauf à un nombre fini de points avec $\|\gamma'\|$ borné dans J ; dans ces conditions, les intégrales sur J dans (6.1) et (6.2) existent, même dans le sens de l'intégration élémentaire de Riemann.

6.1.2 Terminologie

Une fonction continue $\gamma : J \to \mathbb{R}^N$, J étant un intervalle quelconque de $\mathbb{R}$, sera appelée une ***trajectoire*** ; on peut interpréter γ comme la trajectoire d'une particule dans $\mathbb{R}^N$; on dit que γ est ***paramétrée*** par un ***paramètre*** variant dans J ou parcourant J. Dans l'interprétation évoquée, le paramètre est le temps mais on peut donner d'autres interprétations ; par exemple, on peut considérer γ comme décrivant la position d'un fil métallique dans $\mathbb{R}^N$ où le paramètre est la longueur du fil en partant d'un point fixé sur ce dernier.

Une trajectoire γ s'appelle un ***chemin*** si J est compact ; si $J = [a, b]$, on dit que les points $\gamma(a)$, $\gamma(b)$ sont les ***extrémités*** du chemin, $\gamma(a)$ étant son ***origine*** ou ***début*** ou ***point initial*** et $\gamma(b)$ sa ***fin*** ou son ***point final*** ;

on dit aussi que le chemin γ ***relie*** les points $\gamma(a)$ et $\gamma(b)$ ***dans*** $E \subset \mathbb{R}^N$ si $\gamma(J) \subset E$. Le chemin γ s'appelle un ***lacet*** ou ***chemin fermé*** si ses extrémités coïncident, c'est-à-dire si $\gamma(a) = \gamma(b)$ où $\gamma : [a, b] \to \mathbb{R}^N$; le chemin s'appelle ***constant*** ou ***réduit à un point*** si γ est une fonction constante ; on dit que γ est un ***chemin simple*** si l'application γ est injective ; il s'appelle un ***lacet simple*** si c'est un lacet ($\gamma(a) = \gamma(b)$) tel que la restriction de γ à $[a, b[$ est injective. Si $\gamma : [a, b] \to \mathbb{R}^N$ est une application de classe C^1 on dit que γ est un ***chemin*** C^1 ; on dira que γ est un ***chemin*** C^1 ***par morceaux*** s'il existe une subdivision finie de $[a, b]$ donnée par

$$t_0 = a < t_1 < \ldots < t_{n-1} < t_n = b$$

telle que γ est de classe C^1 dans chaque sous-intervalle $[t_{i-1}, t_i]$ et, de plus, γ est continue dans $[a, b]$. Un exemple particulier important d'un chemin C^1 par morceaux est un chemin ***polygonal*** défini dans le paragraphe 1.8.1. Un chemin γ s'appelle ***régulier*** s'il est de classe C^1 et si $\gamma'(t) \neq 0$ pour tout t ; pour un chemin régulier, on a un ***vecteur unitaire tangent*** τ à chaque point du chemin en posant

$$\tau(\gamma(t)) = \frac{\gamma'(t)}{\|\gamma'(t)\|} \,. \tag{6.4}$$

Un chemin $\gamma : [a, b] \to \mathbb{R}^N$ s'appelle ***régulier par morceaux*** s'il est C^1 par morceaux et si l'ensemble des t, où $\gamma'(t)$ existe et est égal à 0, est un ensemble fini. *Les chemins polygonaux sont réguliers par morceaux.*

Une grande partie de la terminologie introduite pour les chemins peut être transférée aux trajectoires ; ainsi, nous pouvons parler des trajectoires de classe C^1 ou de classe C^1 par morceaux, etc.

Soit $\gamma : J \to \mathbb{R}^N$ une trajectoire ; on dira qu'un point x de $\gamma(J)$ est un point sur la trajectoire ou que la trajectoire γ passe par x ; pour un tel point x, l'ensemble

$$E_x = \{t \in J : \gamma(t) = x\}$$

est non vide ; si la cardinalité de E_x est > 1, on dit que x est un point ***multiple*** de γ (de ***multiplicité*** finie n si la cardinalité de E_x est n ; x est un point ***double*** de γ si $n = 2$) ; si E_x contient exactement un point, x sera un point ***simple*** de γ. Un chemin (ou trajectoire) simple γ est tel que tout point x sur γ est un point simple ; pour un lacet $\gamma : [a, b] \to \mathbb{R}^N$, on aura un lacet simple si tout point x sur γ, sauf les extrémités ($\gamma(a) = \gamma(b)$), est un point simple pour γ.

6.1.3 Interprétation intuitive des intégrales curvilignes

Nous reprenons les notations du paragraphe 6.1.1. Si la trajectoire γ représente un fil métallique et $f(\gamma(t))$ est la densité du matériel au point $\gamma(t)$ du fil, alors $f(\gamma(t)) \, \|\gamma'(t)\| \, \mathrm{d}t$ donne la masse de la partie infinitésimale du fil donnée par $\gamma(t)$, t allant de t à $t + \mathrm{d}t$, de sorte que $\int_\gamma f \, \mathrm{d}s$ peut être interprétée comme la masse totale du fil γ. En interprétant f comme une densité de charge, de chaleur ou d'autres grandeurs physiques, on aura d'autres interprétations

de $\int_\gamma f\,\mathrm{d}s$. En prenant $f = 1$, l'intégrale $\int_\gamma \mathrm{d}s$ donne la ***longueur*** de la trajectoire γ :

$$\text{longueur}(\gamma) = \int_\gamma \mathrm{d}s = \int_J \|\gamma'(t)\|\,\mathrm{d}t\,.$$

On peut interpréter ceci comme suit : $\gamma'(t)$ représente le vecteur vitesse au temps t d'une particule ayant la trajectoire γ ; donc, $\|\gamma'(t)\|\,\mathrm{d}t$ représente la distance parcourue par cette particule dans un laps de temps infinitésimal, t à $t + \mathrm{d}t$, d'où une intégration par rapport à t donne la longueur de γ.

Pour une interprétation de $\int_\gamma F\cdot\mathrm{d}\vec{s}$, nous considérons F comme un champ de force ; alors si γ est régulière,

$$F\big(\gamma(t)\big)\cdot\gamma'(t) = \big\{F\big(\gamma(t)\big)\cdot\tau\big(\gamma(t)\big)\big\}\,\|\gamma'(t)\|$$

où τ est le vecteur tangent unitaire défini par (6.4) ; le terme $\{F\cdot\tau\}$ est la composante de la force F dans la direction τ ; donc, $\int_\gamma F\cdot\mathrm{d}\vec{s}$ définie par (6.2) donne le travail accompli par F pendant le parcours γ car elle est la somme des travaux infinitésimaux

$$\{F\cdot\tau\}\,\mathrm{d}s = \{F\cdot\tau\}\,\|\gamma'(t)\|\,\mathrm{d}t\,.$$

Cette discussion montre aussi que l'intégrale (6.2) est un cas particulier de celle de (6.1) si γ est régulier :

$$\int_\gamma F\cdot\mathrm{d}\vec{s} = \int_\gamma (F\cdot\tau)\,\mathrm{d}s\,.$$

Si G est un autre champ vectoriel défini dans $E \supset \gamma(J)$, on peut définir

$$\int_\gamma (F\cdot G)\,\mathrm{d}s$$

d'après (6.1) ; on a vu que $\int_\gamma F\cdot\mathrm{d}\vec{s}$ en est un cas particulier, cela correspond ici à

$$G\big(\gamma(t)\big) = \tau\big(\gamma(t)\big)\,.$$

Nous donnons maintenant une proposition concernant les changements de variables dans les intégrales (6.1) et (6.2) que nous interpréterons par la suite.

6.1.4 Proposition

Soient $\delta : [c\,,d] \to \mathbb{R}^N$ *un chemin* C^1 *par morceaux dans* $E \subset \mathbb{R}^N$, $\theta : [a\,,b] \to [c\,,d]$ *une application strictement monotone de classe* C^1, $\gamma = \delta\circ\theta$, $f : E \to \mathbb{R}$ *une fonction continue,* $F : E \to \mathbb{R}^N$ *un champ vectoriel continu. Alors* $\gamma : [a\,,b] \to \mathbb{R}^N$ *est un chemin* C^1 *par morceaux dans* E *et*

$$\int_\gamma f\,\mathrm{d}s = \int_\delta f\,\mathrm{d}s \tag{6.5}$$

$$\int_\gamma F \cdot \mathrm{d}\vec{s} = \int_\delta F \cdot \mathrm{d}\vec{s} \qquad \textit{si } \theta \textit{ est croissante} \tag{6.6}$$

$$\int_\gamma F \cdot \mathrm{d}\vec{s} = -\int_\delta F \cdot \mathrm{d}\vec{s} \quad \textit{si } \theta \textit{ est décroissante}. \tag{6.7}$$

DÉMONSTRATION

On a $\gamma'(t) = \delta'\big(\theta(t)\big)\,\theta'(t)$ à chaque point t de $[a, b]$ tel que $\theta(t)$ est un point de dérivabilité de δ ; de cette remarque on voit que γ est un chemin C^1 par morceaux avec

$$\gamma\big([a, b]\big) = \delta\big([c, d]\big) \subset E.$$

Aussi,

$$\begin{aligned}\int_\gamma f\,\mathrm{d}s &= \int_a^b f\big(\gamma(t)\big)\,\big\|\gamma'(t)\big\|\,\mathrm{d}t \\ &= \int_a^b f\big(\delta\circ\theta(t)\big)\,\big\|\delta'\big(\theta(t)\big)\big\|\,\big|\theta'(t)\big|\,\mathrm{d}t \\ &= \int_c^d f\big(\delta(u)\big)\,\big\|\delta'(u)\big\|\,\mathrm{d}u = \int_\delta f\,\mathrm{d}s\end{aligned}$$

en faisant la transformation $u = \theta(t)$. De même,

$$\begin{aligned}\int_\gamma F \cdot \mathrm{d}\vec{s} &= \int_a^b F\big(\gamma(t)\big)\cdot\gamma'(t)\,\mathrm{d}t \\ &= \int_a^b \big\{F\big(\delta\circ\theta(t)\big)\cdot\delta'\big(\theta(t)\big)\big\}\,\theta'(t)\,\mathrm{d}t \\ &= \pm\int_c^d F\big(\delta(u)\big)\cdot\delta'(u)\,\mathrm{d}u = \pm\int_\delta F\cdot d\vec{s}\end{aligned}$$

avec le signe + si $\theta'(t) \geqslant 0$ et avec le signe − si $\theta'(t) \leqslant 0$, selon l'explication donnée dans le paragraphe 5.7.1. ♦

6.1.5 Explications au sujet de la proposition 6.1.4

Si θ de la proposition 6.1.4 est telle que $\theta'(t) \neq 0$ pour tout t dans $[a, b]$ alors θ est un C^1-difféomorphisme entre $[a, b]$ et $[c, d]$; dans ce cas, ou bien $\theta'(t) > 0$ pour tout t ou bien $\theta'(t) < 0$ pour tout t, t variant dans $[a, b]$; on dit alors que le chemin $\gamma = \delta \circ \theta$ est obtenu du chemin δ par un ***changement de paramètre*** ou par un ***reparamétrage***. Le reparamétrage est ***direct*** si $\theta' > 0$, ***opposé*** si $\theta' < 0$; si $\gamma = \delta \circ \theta$ et θ est un reparamétrage direct, on dit alors que γ est ***équivalent*** à δ ; si θ est un reparamétrage opposé, on dit alors que γ est ***opposé*** à δ. Si γ est équivalent ou opposé à δ, alors,

$$\gamma\big([a, b]\big) = \delta\big([c, d]\big).$$

L'ensemble $\gamma([a\,,b]) \subset \mathbb{R}^N$ s'appelle la ***trace*** du chemin γ ; ainsi, si γ et δ sont deux chemins équivalents ou opposés, leurs traces sont identiques. Si γ et δ sont équivalents, leurs débuts et fins sont identiques. Si γ et δ sont opposés, le début de γ est la fin de δ et la fin de γ est le début de δ ; dans ce cas, on dit aussi que les sens de parcours de γ et δ sont opposés.

REMARQUES

- Un chemin $\gamma : [a\,,b] \to \mathbb{R}^N$ est équivalent à lui-même.
- Si $\gamma : [a\,,b] \to \mathbb{R}^N$, $\delta : [c\,,d] \to \mathbb{R}^N$ sont deux chemins, si γ est équivalent à δ alors δ est équivalent à γ ; en effet, $\gamma = \delta \circ \theta$, $\theta : [a\,,b] \to [c\,,d]$ un C^1-difféomorphisme ; alors $\delta = \gamma \circ \theta^{-1}$, $\theta^{-1} : [c\,,d] \to [a\,,b]$.
- Si γ est équivalent à δ et δ est équivalent à η alors γ est équivalent à η, où γ, δ, η sont trois chemins de $\mathbb{R}^N$ (laissé en exercice).
- Tout chemin γ est équivalent à un chemin $\delta : [0\,,1] \to \mathbb{R}^N$ ou à un chemin $\delta : [0\,,2\pi] \to \mathbb{R}^N$. En effet, si $\gamma : [a\,,b] \to \mathbb{R}^N$, il suffit de prendre

$$\theta(t) = a + (b-a)t\,, \quad 0 \leqslant t \leqslant 1\,,$$

  et de poser $\delta = \gamma \circ \theta$ pour avoir un chemin $\delta : [0\,,1] \to \mathbb{R}^N$ équivalent à γ ; de même, pour avoir $\delta : [0\,,2\pi] \to \mathbb{R}^N$ équivalent à γ, on prendra $\delta = \gamma \circ \theta$ avec

$$\theta(t) = a + \frac{b-a}{2\pi}\,t\,, \quad 0 \leqslant t \leqslant 2\pi\,.$$

- Si $\gamma : [a\,,b] \to \mathbb{R}^N$ est un chemin, alors $\widetilde{\gamma}(t) = \gamma(a+b-t)$ définit un chemin $\widetilde{\gamma} : [a\,,b] \to \mathbb{R}^N$ qui est opposé à γ (laissé en exercice).

Avec ces terminologies, on peut reformuler la proposition 6.1.4 comme suit (les hypothèses de régularités sur γ, δ, f, F étant maintenues) :

- $\displaystyle\int_\gamma f\,\mathrm{d}s = \int_\delta f\,\mathrm{d}s$ si le chemin γ (δ) est obtenu à partir de δ (γ) par un reparamétrage ;
- $\displaystyle\int_\gamma F\cdot\mathrm{d}\vec{s} = \int_\delta F\cdot\mathrm{d}\vec{s}$ si γ, δ sont deux chemins équivalents ;
- $\displaystyle\int_\gamma F\cdot\mathrm{d}\vec{s} = -\int_\delta F\cdot\mathrm{d}\vec{s}$ si γ, δ sont deux chemins opposés ;

Il est clair que si δ est un lacet, le chemin γ équivalent ou opposé à δ sera aussi un lacet.

6.1.6 Représentation d'un lacet

Il est commode d'utiliser une autre façon de représenter un lacet dans $\mathbb{R}^N$. Soit C le cercle unité dans $\mathbb{R}^2$:

$$C = \{(u,v) : u^2 + v^2 = 1\}\,;$$

soit $\varphi : C \to \mathbb{R}^N$ une fonction continue ; alors,

$$\gamma(t) = \varphi(\cos t, \sin t), \quad 0 \leqslant t \leqslant 2\pi, \tag{6.8}$$

définit un lacet $\gamma : [0, 2\pi] \to \mathbb{R}^N$. D'autre part, si $\gamma : [0, 2\pi] \to \mathbb{R}^N$ est un lacet, *on peut définir une fonction continue* $\varphi : C \to \mathbb{R}^N$ *telle que (6.8) est vérifiée* ; en effet, si $(u, v) \in C$, il existe un unique $t \in [0, 2\pi[$ tel que

$$(u, v) = (\cos t, \sin t) ;$$

alors, si

$$\varphi(u, v) = \gamma(t), \quad 0 \leqslant t < 2\pi,$$

on aura $\varphi : C \to \mathbb{R}^N$, *une fonction continue, telle que (6.8) est vérifiée*. Notons que la correspondance $(u, v) \mapsto t$ n'est pas continue comme une application de C sur $[0, 2\pi[$; néanmoins, φ est continue sur C car $\gamma(0) = \gamma(2\pi)$. Puisque tout lacet est équivalent à un lacet défini dans $[0, 2\pi]$, on voit qu'un lacet dans $\mathbb{R}^N$ peut toujours être considéré comme une application continue $\varphi : C \to \mathbb{R}^N$. Ainsi, un lacet simple peut être représenté comme une application continue injective $\varphi : C \to \mathbb{R}^N$.

6.1.7 Paramétrage naturel

Soit $\gamma : [a, b] \to \mathbb{R}^N$ un chemin ; nous discutons dans ce paragraphe la notion de la « longueur » de γ d'un point de vue général qui permet de mieux apprécier la formule donnée pour la longueur dans le paragraphe 6.1.3. Soit π une subdivision de l'intervalle $[a, b]$ par des points $a = t_0 < t_1 < \ldots < t_n = b$; à chaque subdivision π, on peut attacher le nombre $\ell_\gamma(\pi)$, la longueur du chemin polygonal formé par les segments successifs $[\gamma(t_{i-1}), \gamma(t_i)]$, $i = 1, \ldots, n$; autrement dit,

$$\ell_\gamma(\pi) = \sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\| .$$

On *définit*

$$\text{longueur}(\gamma) = \sup \ell_\gamma(\pi),$$

le supremum étant pris sur toutes les subdivisions π de $[a, b]$; on dit que le chemin est ***rectifiable*** si longueur$(\gamma) < \infty$. *Une condition nécessaire et suffissante pour que* γ *soit rectifiable est que, si* $\gamma = (\gamma_1, \ldots, \gamma_N)$, *chacune des fonctions réelles* γ_j *soit une fonction à* ***variation bornée*** *dans le sens que*

$$\sup_{\pi} \sum_{i=1}^{n} |\gamma_j(t_i) - \gamma_j(t_{i-1})| < \infty .$$

La preuve de cette affirmation est immédiate grâce à l'inégalité élémentaire suivante :

$$\max_j |\gamma_j(t_i) - \gamma_j(t_{i-1})| \leqslant \|\gamma(t_i) - \gamma(t_{i-1})\| \leqslant \sqrt{N} \max_j |\gamma_j(t_i) - \gamma_j(t_{i-1})| .$$

Si γ est C^1 par morceaux, il n'est pas difficile de démontrer qu'alors γ est rectifiable et longueur(γ), définie ci-dessus, est donnée par la formule du paragraphe 6.1.3 :

$$\text{longueur}(\gamma) = \int_a^b \|\gamma'(t)\| \, dt .$$

Un des grands succès de la théorie de Lebesgue est la démonstration du fait que *si γ est rectifiable alors $\gamma'(t)$ existe p.p. dans* $[a, b]$; la preuve de cette dernière affirmation est subtile et nous ne la donnons pas ici ; plus généralement, on a l'inégalité suivante *pour tout chemin rectifiable* γ :

$$\text{longueur}(\gamma) \geqslant \int_a^b \|\gamma'(t)\| \, dt .$$

Il y a égalité dans la dernière inégalité si et seulement si les fonctions γ_i sont *absolument continues*, une notion que nous ne développerons pas ici (cf. Rudin [37, p. 145]) ; en particulier, si les γ_i sont C^1 par morceaux, il y a égalité.

Soit $\gamma : [a, b] \to \mathbb{R}^N$ un chemin rectifiable : on dit que le paramétrage de γ est ***naturel*** si, pour $a \leqslant t_1 < t_2 \leqslant b$,

$$\text{longueur}\left(\gamma|_{[t_1, t_2]}\right) = t_2 - t_1 .$$

Nous démontrons que *si γ est un chemin régulier* (§ 6.1.2) *alors γ est équivalent à un chemin $\delta : [0, L] \to \mathbb{R}^N$ ayant un paramétrage naturel.* En effet, posons

$$\lambda(t) = \int_a^t \|\gamma'(u)\| \, du , \quad a \leqslant t \leqslant b ;$$

alors $\lambda(a) = 0$, $\lambda(b) = L$, la longueur de γ et $\lambda : [a, b] \to [0, L]$ est une fonction strictement croissante de classe C^1 avec

$$\lambda'(t) = \|\gamma'(t)\| > 0 , \quad a \leqslant t \leqslant b .$$

Soit $\theta : [0, L] \to [a, b]$ la fonction inverse de λ ; ainsi, $\lambda\big(\theta(s)\big) = s$, $0 \leqslant s \leqslant L$, d'où

$$\lambda'\big(\theta(s)\big)\, \theta'(s) = 1 , \quad 0 \leqslant s \leqslant L .$$

Evidemment, θ est un C^1-difféomorphisme entre $[0, L]$ et $[a, b]$; donc $\delta = \gamma \circ \theta$ est un chemin équivalent à γ ; de plus, $\|\delta'(s)\| = 1$ pour $0 \leqslant s \leqslant L$, car

$$\|\delta'(s)\| = \|\gamma'\big(\theta(s)\big)\| \, \theta'(s) = \lambda'\big(\theta(s)\big)\, \theta'(s) = 1 ;$$

ainsi,

$$\text{longueur}\left(\delta|_{[s_1, s_2]}\right) = \int_{s_1}^{s_2} ds = s_2 - s_1$$

de sorte que le paramètre de δ est naturel. Ceci établit notre affirmation.

La proposition suivante établit une estimation utile pour les intégrales curvilignes.

6.1.8 Proposition

Soient $\gamma : [a, b] \to \mathbb{R}^N$ *un chemin* C^1 *par morceaux,* $E = \gamma([a, b])$, $f : E \to \mathbb{R}$ *et* $F : E \to \mathbb{R}^N$ *continues ; soit* L *la longueur de* γ.

(i) Si $M = \sup_{x \in E} |f(x)|$ *alors*

$$\left| \int_\gamma f \, ds \right| \leqslant LM .$$

(ii) Si $M = \sup_{x \in E} \|F(x)\|$ *alors*

$$\left| \int_\gamma F \cdot d\vec{s} \right| \leqslant LM .$$

DÉMONSTRATION

$$\text{(i)} \quad \left| \int_\gamma f \, ds \right| = \left| \int_a^b f(\gamma(t)) \|\gamma'(t)\| \, dt \right|$$

$$\leqslant \int_a^b |f(\gamma(t))| \, \|\gamma'(t)\| \, dt \leqslant M \int_a^b \|\gamma'(t)\| \, dt = ML .$$

$$\text{(ii)} \quad \left| \int_\gamma F \cdot d\vec{s} \right| = \left| \int_a^b F(\gamma(t)) \cdot \gamma'(t) \, dt \right|$$

$$\leqslant \int_a^b |F(\gamma(t)) \cdot \gamma'(t)| \, dt$$

$$\leqslant \int_a^b \|F(\gamma(t))\| \, \|\gamma'(t)\| \, dt \leqslant M \int_a^b \|\gamma'(t)\| \, dt = ML$$

où l'on a utilisé l'inégalité de Cauchy-Schwarz : $|x \cdot y| \leqslant \|x\| \, \|y\|$. ♦

6.1.9 Champ dérivant d'un potentiel

Supposons que $F : \Omega \to \mathbb{R}^N$ est un champ vectoriel défini dans un ensemble ouvert Ω de $\mathbb{R}^N$; Si $F = \operatorname{grad} u$ où $u \in C^1(\Omega)$, nous dirons que F est un champ ***dérivant d'un potentiel*** u et que u est un ***potentiel scalaire*** ou simplement un ***potentiel*** pour le champ F. *Si* Ω *est un domaine, le potentiel* u *de* F, *s'il existe, est unique à une constante additive près.* En effet, si $\operatorname{grad} u = \operatorname{grad} v = F$ dans Ω, on aura $\operatorname{grad}(u - v) = 0$ d'où $u - v =$ constante dans Ω (sect. 2.3). Une propriété importante d'un champ F dérivant d'un potentiel $u \in C^1(\Omega)$, Ω ouvert dans $\mathbb{R}^N$, est la suivante : *si* γ *est un chemin* C^1 *par morceaux dans* Ω *et* $F = \operatorname{grad} u$ *alors*

$$\int_\gamma F \cdot d\vec{s} = u(B) - u(A) \tag{6.9}$$

où A *est le début de* γ *et* B *la fin de* γ.

En effet, si $\gamma : [a, b] \to \mathbb{R}^N$ est de classe C^1 et $A = \gamma(a)$, $B = \gamma(b)$,

$$\int_\gamma F \cdot d\vec{s} = \int_a^b \nabla u\big(\gamma(t)\big) \cdot \gamma'(t)\, dt$$
$$= \int_a^b \frac{d}{dt} \big\{u\big(\gamma(t)\big)\big\}\, dt = u\big(\gamma(b)\big) - u\big(\gamma(a)\big) = u(B) - u(A)\,.$$

Ceci démontre (6.9) dans le cas d'un chemin C^1 ; si $\gamma : [a, b] \to \mathbb{R}^N$ est C^1 par morceaux, on aura $a = t_0 < t_1 < \ldots < t_n = b$ tels que $\gamma\big|_{[t_{i-1}, t_i]}$ de classe C^1, $i = 1, \ldots, n$; donc, du calcul qui précède, on déduit que

$$\int_\gamma F \cdot d\vec{s} = \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{d}{dt} \big\{u\big(\gamma(t)\big)\big\}\, dt$$
$$= \sum_{i=1}^n \big\{u\big(\gamma(t_i)\big) - u\big(\gamma(t_{i-1})\big)\big\}$$
$$= u\big(\gamma(b)\big) - u\big(\gamma(a)\big) = u(B) - u(A)\ ;$$

ceci établit (6.9) complètement.

Un corollaire immédiat de (6.9) est le suivant : *si* $F = \nabla u$, $u \in C^1(\Omega)$, Ω *ouvert dans* $\mathbb{R}^N$ *et si* γ_1, γ_2 *sont deux chemins,* C^1 *par morceaux, dans* Ω *ayant les mêmes débuts et fins, alors,*

$$\int_{\gamma_1} F \cdot d\vec{s} = \int_{\gamma_2} F \cdot d\vec{s}\,. \tag{6.10}$$

En particulier, si γ *est un lacet,* C^1 *par morceaux, dans* Ω,

$$\int_\gamma F \cdot d\vec{s} = 0\,. \tag{6.11}$$

On verra par la suite que les propriétés d'un champ vectoriel F exprimées par (6.9), (6.10) et (6.11) sont équivalentes (prop. 6.1.11).

Si $F = (F_1, \ldots, F_N)$ est un champ vectoriel de classe C^1 dans une partie ouverte Ω de $\mathbb{R}^N$, alors le théorème sur l'égalité des dérivées mixtes (sect. 3.2) donne une condition *nécessaire* pour que F dérive d'un potentiel u : *il faut que*

$$D_j F_i = D_i F_j\,, \quad 1 \leqslant i, j \leqslant N\,, \tag{6.12}$$

En effet, si $F \in C^1(\Omega; \mathbb{R}^N)$ et $F = \nabla u$ dans Ω, on a que $u \in C^2(\Omega)$ et $F_i = D_i u$ dans Ω ; (6.12) est une conséquence immédiate du fait que

$$D_j D_i u = D_i D_j u\,, \quad 1 \leqslant i, j \leqslant N\,.$$

Noter que *pour* $N = 3$, *(6.12) équivaut à la condition* $\operatorname{rot} F = 0$. Par la suite, on verra que si Ω satisfait une condition géométrique appropriée, un champ vectoriel $F \in C^1(\Omega; \mathbb{R}^N)$ qui remplit la condition (6.12) dérive nécessairement d'un potentiel u (prop. 6.1.15).

Pour la commodité de l'exposé, nous introduisons quelques notations et conventions supplémentaires.

6.1.10 Notations et conventions

Si α et β sont deux chemins tels que l'origine de β est la fin de α, on notera par $\alpha \vee \beta$ le chemin formé par la ***juxtaposition*** de β après α ; analytiquement, si

$$\alpha : [a, b] \to \mathbb{R}^N , \quad \beta : [c, d] \to \mathbb{R}^N , \quad \alpha(b) = \beta(c)$$

alors

$$(\alpha \vee \beta)(t) = \begin{cases} \alpha(t) & \text{si } a \leqslant t \leqslant b \\ \beta(t + c - b) & \text{si } b \leqslant t \leqslant b + d - c \end{cases}$$

dont le début est $\alpha(a)$ et la fin $\beta(d)$. Plus généralement, par $\alpha_1 \vee \cdots \vee \alpha_n$, on représente un chemin obtenu par la juxtaposition successive des chemin $\alpha_2, \ldots, \alpha_n$ qui sont tels que l'origine de α_j coïncide avec la fin de α_{j-1}, $2 \leqslant j \leqslant n$. Il est clair que si les α_j sont C^1 par morceaux, c'est aussi le cas de $\alpha_1 \vee \cdots \vee \alpha_n$ et

$$\int_{\alpha_1 \vee \cdots \vee \alpha_n} f \, \mathrm{d}s = \sum_{j-1}^{n} \int_{\alpha_j} f \, \mathrm{d}s$$

$$\int_{\alpha_1 \vee \cdots \vee \alpha_n} F \cdot \mathrm{d}\vec{s} = \sum_{j-1}^{n} \int_{\alpha_j} F \cdot \mathrm{d}\vec{s}$$

pourvu que les intégrales concernées existent. On utilisera cette écriture et convention aussi lorsque les chemins $\alpha_1, \ldots, \alpha_n$ sont « deux à deux disjoints », c'est-à-dire que leurs traces sont deux à deux disjointes.

Il sera par la suite commode de décrire certains chemins comme la juxtaposition successives des autres. Aussi, si C est un segment ou un arc d'un cercle ou une courbe simple bien connue de la géométrie élémentaire dont les extrémités sont x et y, on utilisera les descriptions du type suivant pour définir un chemin α reliant x à y : α est un chemin formé de C allant de x à y. Analytiquement, ceci veut dire que $\alpha : [a, b] \to \mathbb{R}^N$ est un chemin simple C^1 tel que $\alpha(a) = x$, $\alpha(b) = y$ et que la trace de α est C ; puisque la formule exacte spécifiant α ne joue aucun rôle dans les calculs théoriques (prop. 6.1.4), cette description de α sera généralement suffisante.

Pour la proposition 6.1.11, il est utile d'introduire la terminologie suivante : soit F un champ vectoriel dans une partie ouverte Ω de $\mathbb{R}^N$; on dira que F est un champ ***conservatif*** si l'égalité (6.10) est valable pour tous les chemins γ_1, γ_2 dans Ω, C^1 par morceaux et ayant les mêmes débuts et fins.

6.1.11 Proposition

Soit F un champ vectoriel continu dans une partie ouverte non vide Ω de $\mathbb{R}^N$. Les affirmations suivantes sont équivalentes :

(i) F dérive d'un potentiel $u \in C^1(\Omega)$;
(ii) F est un champ conservatif ;
(iii) $\int_\gamma F \cdot d\vec{s} = 0$ pour tout lacet γ, C^1 par morceaux, dans Ω.

DÉMONSTRATION

(i)⇒(ii) Ceci a été vu dans (6.10) au paragraphe 6.1.9.

(ii)⇒(iii) Soit $\gamma : [a, b] \to \Omega$ un lacet C^1 par morceaux. Si γ est un chemin constant, $\int_\gamma F \cdot d\vec{s}$ est évidemment 0 ; si non, prenons un c tel que $a < c < b$ avec $\gamma(c) \neq \gamma(a)$; posons α pour le chemin $\gamma|_{[a,c]}$ et β pour le chemin $\gamma|_{[c,b]}$; alors $\gamma = \alpha \vee \beta$. Si $\widetilde{\beta}$ est un chemin opposé à β (§ 6.1.5), les deux chemins α et $\widetilde{\beta}$ auront le même début $\gamma(a)$ et la même fin $\gamma(c)$; de (ii) on aura

$$\int_\alpha F \cdot d\vec{s} = \int_{\widetilde{\beta}} F \cdot d\vec{s}$$

d'où

$$\int_\gamma F \cdot d\vec{s} = \int_\alpha F \cdot d\vec{s} + \int_\beta F \cdot d\vec{s} = \int_\alpha F \cdot d\vec{s} - \int_{\widetilde{\beta}} F \cdot d\vec{s} = 0 .$$

(iii)⇒(ii) Si γ_1, γ_2 sont deux chemins, C^1 par morceaux, dans Ω, ayant les mêmes débuts et fins, le chemin $\gamma = \gamma_1 \vee \widetilde{\gamma}_2$, $\widetilde{\gamma}_2$ étant un chemin opposé à γ_2, sera un lacet, C^1 par morceaux, dans Ω ; de (iii),

$$\int_\gamma F \cdot d\vec{s} = 0$$

d'où

$$0 = \int_\gamma F \cdot d\vec{s} = \int_{\gamma_1} F \cdot d\vec{s} + \int_{\widetilde{\gamma}_2} F \cdot d\vec{s} = \int_{\gamma_1} F \cdot d\vec{s} - \int_{\gamma_2} F \cdot d\vec{s},$$

ce qui établit (ii).

(ii)⇒(i) On sait que tout ensemble ouvert non vide dans $\mathbb{R}^N$ est une réunion des domaines deux à deux disjoints (§ 1.8.3 (6)) ; il suffit donc d'établir (i) sous l'hypothèse que Ω est un domaine (c'est-à-dire ouvert, connexe, non vide). Fixons un point $p \in \Omega$; pour tout $x \in \Omega$, il existe un chemin γ_x dans Ω reliant p à x ; γ_x peut être choisi comme un chemin polygonal (même simple ; prop. 1.8.2(iv)), en particulier, C^1 par morceaux. Posons

$$u(x) = \int_{\gamma_x} F \cdot d\vec{s}.$$

Selon (ii), $u(x)$ ne dépend pas du choix particulier de γ_x ; vérifions que $\nabla u = F$ dans Ω, c'est-à-dire

$$\mathrm{D}_i u(x) = F_i(x)\,, \quad 1 \leqslant i \leqslant N\,,\ x \in \Omega\,,$$

si $F = (F_1, \ldots, F_N)$. Nous ferons cette vérification pour $i = 1$. Soit $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^N$; pour $\varepsilon \in \mathbb{R}^*$, soit γ_x^ε le chemin obtenu par la juxtaposition de γ_x et du segment $[\,x\,,\,x + \varepsilon e_1\,]$; si ε est petit, γ_x^ε sera un chemin dans Ω et

$$\begin{aligned} u(x + \varepsilon e_1) &= \int_{\gamma_x^\varepsilon} F \cdot \mathrm{d}\vec{s} = \int_{\gamma_x} F \cdot \mathrm{d}\vec{s} + \int_{[\,x\,,\,x+\varepsilon e_1\,]} F \cdot \mathrm{d}\vec{s} \\ &= u(x) + \int_{[\,x\,,\,x+\varepsilon e_1\,]} F \cdot \mathrm{d}\vec{s} \end{aligned}$$

d'où

$$\frac{1}{\varepsilon}\left\{u(x + \varepsilon e_1) - u(x)\right\} = \frac{1}{\varepsilon}\int_{[\,x\,,\,x+\varepsilon e_1\,]} F \cdot \mathrm{d}\vec{s} = \int_0^1 F_1(x + t\varepsilon e_1)\,\mathrm{d}t$$

en paramétrant le segment $[\,x\,,\,x + \varepsilon e_1\,]$ par

$$\delta(t) = x + t\varepsilon e_1\,, \quad 0 \leqslant t \leqslant 1\,,\ \delta'(t) = \varepsilon e_1\,.$$

Puisque

$$\lim_{\varepsilon \to 0} \int_0^1 F_1(x + t\varepsilon e_1)\,\mathrm{d}t = \int_0^1 \lim_{\varepsilon \to 0} F_1(x + t\varepsilon e_1)\,\mathrm{d}t = F_1(x)$$

à cause de la continuité de F_1, on voit que

$$\mathrm{D}_1 u(x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left\{u(x + \varepsilon e_1) - u(x)\right\} = F_1(x)\,, \quad x \in \Omega\,. \qquad \blacklozenge$$

REMARQUE. La justification pour la relation

$$\lim_{\varepsilon \to 0} \int_0^1 F_1(x + t\varepsilon e_1)\,\mathrm{d}t = \int_0^1 \lim_{\varepsilon \to 0} F_1(x + t\varepsilon e_1)\,\mathrm{d}t$$

est immédiate à partir du théorème de convergence dominée (prop. 5.2.8(iii)) ou de la proposition 5.3.1 car il existe M et ε_0 tels que $\left|F_1(x + t\varepsilon e_1)\right| \leqslant M$ pour $0 \leqslant t \leqslant 1$ si $|\varepsilon| \leqslant \varepsilon_0$ à cause de la continuité de F_1 ; mais l'on peut aussi justifier le résultat

$$\lim_{\varepsilon \to 0} \int_0^1 F_1(x + t\varepsilon e_1)\,\mathrm{d}t = F_1(x)$$

directement à partir de la relation

$$\left|\int_0^1 F_1(x+t\varepsilon e_1)\,\mathrm{d}t - F_1(x)\right| \leqslant \int_0^1 \left|F_1(x+t\varepsilon e_1) - F_1(x)\right| \mathrm{d}t < \delta$$

valable si $|\varepsilon| < \varepsilon_0$; il suffit de choisir ε_0 tel que $\left|F_1(y) - F_1(x)\right| < \delta$ dès que $|y-x| < \varepsilon_0$, $y \in \Omega$, ce qui est possible à cause de la continuité de F_1 au point x.

6.1.12 Compléments à la proposition 6.1.11

Il est utile d'ajouter aux trois affirmations équivalentes contenues dans la proposition 6.1.11, les deux autres qui suivent :

(ii′) (6.10) est vérifié pour les chemins polygonaux simples γ_1, γ_2 dans Ω ayant les mêmes débuts et fins ;

(iii′) $\int_\gamma F \cdot \mathrm{d}\vec{s} = 0$ *pour tout lacet polygonal simple γ dans Ω.*

La démonstration que (ii′) et (iii′) sont équivalentes aux trois affirmations de la proposition 6.1.11 peut être faite comme suit : établir que (ii′) et (iii′) sont équivalentes et voir que (ii′)⇒(i). Les preuves de

$$\text{(ii′)} \Longrightarrow \text{(i)}, \quad \text{(ii′)} \Longrightarrow \text{(iii′)}$$

sont exactement comme celles pour

$$\text{(ii)} \Longrightarrow \text{(i)}, \quad \text{(ii)} \Longrightarrow \text{(iii)}$$

données ci-dessus ; il suffit donc de voir que (iii′)⇒(ii′). Soient γ_1, γ_2 deux chemins polygonaux simples dans Ω ayant les mêmes débuts et fins. Si γ_1, γ_2 ne se croisent nulle part sauf au début et à la fin, alors $\gamma_1 \vee \widetilde{\gamma}_2 = \gamma$ sera un lacet polygonal simple dans Ω (où $\widetilde{\gamma}_2$ est un chemin opposé à γ_2) ; alors,

$$0 = \int_\gamma F \cdot \mathrm{d}\vec{s} = \int_{\gamma_1} F \cdot \mathrm{d}\vec{s} + \int_{\widetilde{\gamma}_2} F \cdot \mathrm{d}\vec{s} = \int_{\gamma_1} F \cdot \mathrm{d}\vec{s} - \int_{\gamma_2} F \cdot \mathrm{d}\vec{s}.$$

Le cas général peut être traité par induction sur le nombre m (fini) de croisements et le nombre n (fini) de segments de γ_1, γ_2 qui se superposent. Nous ne donnons pas les détails fastidieux d'un tel raisonnement.

6.1.13 Explication physique du terme conservatif

Considérons le mouvement d'une particule de masse $m > 0$ dans un domaine Ω de $\mathbb{R}^3$ sous l'influence d'un champ de force F ; nous supposons que $F \in C(\Omega; \mathbb{R}^3)$. Si $x(t)$ est la position de la particule au temps t, le modèle mathématique usuel (la seconde loi de Newton) donne alors

$$mx''(t) = F\big(x(t)\big)\,.$$

Pour $t_0 < t_1$, soit γ le chemin donné par la trajectoire de la particule durant l'intervalle de temps $[\,t_0\,,\,t_1\,]$, c'est-à-dire

$$\gamma(t) = x(t)\,, \quad t_0 \leqslant t \leqslant t_1\,.$$

Alors,

$$\int_\gamma F \cdot \mathrm{d}\vec{s} = \int_{t_0}^{t_1} F\big(x(t)\big) \cdot x'(t)\,\mathrm{d}t = \int_{t_0}^{t_1} mx''(t) \cdot x'(t)\,\mathrm{d}t$$
$$= \int_{t_0}^{t_1} \frac{1}{2}\, m \frac{\mathrm{d}}{\mathrm{d}t}\, v^2(t)\,\mathrm{d}t = \frac{1}{2}\, mv^2(t_1) - \frac{1}{2}\, mv^2(t_0)$$

où

$$v^2(t) = \big\|x'(t)\big\|^2 = x'(t) \cdot x'(t)\,, \quad \frac{\mathrm{d}}{\mathrm{d}t}\, v^2(t) = 2x''(t) \cdot x'(t)\,.$$

L'expression

$$K(t) = \frac{1}{2}\, mv^2(t)$$

s'appelle l'***énergie cinétique*** de la particule au temps t ; en interprétant $\int_\gamma F \cdot \mathrm{d}\vec{s}$ comme le travail accompli pendant $[\,t_0\,,\,t_1\,]$, le calcul précédent montre que ce travail est égal au changement dans l'énergie cinétique $K(t_1) - K(t_0)$. Si F est un champ conservatif, on a $F = \nabla u$, pour une fonction $u \in C^1(\Omega)$, selon la proposition 6.1.11 et

$$\int_\gamma F \cdot \mathrm{d}\vec{s} = u\big(x(t_1)\big) - u\big(x(t_0)\big)$$

d'où, en posant $-U(t) = u\big(x(t)\big)$, on aura

$$K(t_1) - K(t_0) = -U(t_1) + U(t_0)\,.$$

Ceci veut dire que

$$E(t) = U(t) + K(t)$$

est une *constante* pour tout t ; en interprétant $U(t)$ comme l'***énergie potentielle*** de la particule au temps t, on interprète la constance de $E(t)$ comme la ***loi de la conservation d'énergie totale***. Un système *conservatif* comme celui formé d'une particule dans un champ conservatif préserve son énergie.

La discussion précédente peut être généralisée à un système de n particules.

6.1.14 Ensemble étoilé

Un ensemble $E \subset \mathbb{R}^N$ se dit ***étoilé*** s'il existe un point $p \in E$ tel que pour tout $x \in E$, le segment $[\,p\,,\,x\,]$ est contenu dans E ; on dit alors que E est étoilé par rapport au point p. Un ensemble convexe E est étoilé par rapport à tout point de E. Un ensemble étoilé E est évidemment connexe par arcs car,

si x, y sont dans E et si E est étoilé par rapport au point $p \in E$, le chemin polygonal

$$[x, p] \vee [p, y]$$

relie x à y dans E ; ainsi, un ensemble étoilé est connexe.

Il est clair qu'il y a beaucoup d'ensembles étoilés non convexes.

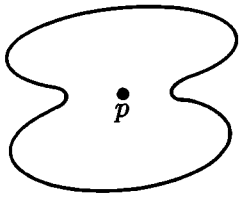


Fig. 6.1 Ensemble non convexe étoilé.

Un exemple simple d'un ensemble non étoilé est $E = \mathbb{R}^N \setminus \{0\}$; en effet, si $p \in E$, le segment $[p, -p]$ doit passer par 0 de sorte que E n'est étoilé par rapport à aucun de ces points. Noter que, pour $N \geqslant 2$, E est un domaine.

La proposition suivante concerne les champs vectoriels dans les domaines étoilés.

6.1.15 Proposition

Soit F un champ vectoriel continu dans un domaine étoilé Ω de $\mathbb{R}^N$.

(i) Supposons que Ω est étoilé par rapport au point $p \in \Omega$; alors F est conservatif si et seulement si

$$\int_\gamma F \cdot \mathrm{d}\vec{s} = 0$$

pour tout lacet triangulaire γ dans Ω ayant p comme un de ses trois sommets.

(ii) Supposons que $F = (F_1, \ldots, F_N)$ est de classe C^1 ; alors F est conservatif si et seulement si

$$\mathrm{D}_i F_j = \mathrm{D}_j F_i\,, \quad 1 \leqslant i, j \leqslant N\,.$$

DÉMONSTRATION

(i) Posons

$$u(x) = \int_{[p,x]} F \cdot d\vec{s}, \quad x \in \Omega,$$

alors $D_i u = F_i$, $i = 1, \ldots, N$, dans Ω ; vérifions ceci pour $i = 1$. Pour ε réel et petit, $x + \varepsilon e_1 \in \Omega$, $e_1 = (1, 0, \ldots, 0)$; on vérifie d'abord que

$$u(x + \varepsilon e_1) - u(x) = \int_{[x, x+\varepsilon e_1]} F \cdot d\vec{s}\,; \tag{6.13}$$

ceci est évident si p, x et $x + \varepsilon e_1$ sont collinéaires ; sinon, ces trois points forment un triangle donnant un lacet triangulaire γ dans Ω avec

$$\int_\gamma F \cdot d\vec{s} = 0$$

ce qui donne

$$\int_{[p,x]} F \cdot d\vec{s} + \int_{[x, x+\varepsilon e_1]} F \cdot d\vec{s} + \int_{[x+\varepsilon e_1, p]} F \cdot d\vec{s} = 0$$

d'où la formule (6.13). De (6.13), on démontre que $D_1 u(x) = F_1(x)$, exactement comme dans la démonstration de (ii)⇒(i) de la proposition 6.1.11. D'autre part, si u est conservatif, $\int_\gamma F \cdot d\vec{s} = 0$ à cause de (iii) de la proposition 6.1.11.

(ii) La nécessité des conditions $D_i F_j = D_j F_i$ était déjà vue dans (6.12) du paragraphe 6.1.9. Nous démontrons maintenant que ces conditions entraînent que F est de la forme grad u, ce qui équivaut à la condition que F est conservatif. Supposons d'abord que Ω est étoilé par rapport au point $0 \in \Omega$; posons

$$u(x) = \int_0^1 F(tx) \cdot x \, dt, \quad x \in \Omega,$$

et vérifions que $D_i u(x) = F_i(x)$, $1 \leqslant i \leqslant N$, ce que nous ferons pour $i = 1$. Du fait que

$$F(tx) \cdot x = F_1(tx)\, x_1 + \cdots + F_N(tx)\, x_N,$$

on obtient, après un petit calcul, que

$$\begin{aligned} \frac{\partial}{\partial x_1}\{F(tx) \cdot x\} &= \sum_{i=1}^N D_1 F_i(tx) t x_i + F_1(tx) \\ &= \sum_{i=1}^N D_i F_1(tx) t x_i + F_1(tx) \\ &= \frac{d}{dt} F_1(tx) + F_1(tx) = \frac{d}{dt}\{t F_1(tx)\} \end{aligned}$$

en utilisant $\mathrm{D}_1 F_i = \mathrm{D}_i F_1$. Ainsi,

$$\begin{aligned}\mathrm{D}_1 u(x) &= \int_0^1 \frac{\partial}{\partial x_1}\{F(tx)\cdot x\}\,\mathrm{d}t \\ &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\{tF_1(tx)\}\,\mathrm{d}t = F_1(x)\end{aligned}$$

où l'on a utilisé la possibilité de la dérivation sous le signe intégrale $\int_0^1$ grâce au corollaire 5.3.4.

Supposons maintenant que Ω soit étoilé par rapport au point $p \in \Omega$; posons alors

$$u(x) = \int_0^1 F\big(p + t(x-p)\big)\cdot(x-p)\,\mathrm{d}t\,, \quad x \in \Omega\,.$$

On vérifie du cas précédent que $\operatorname{grad} u = F$ dans Ω. En effet, en posant $y = x - p$ et

$$G(y) = F(p+y)\,, \quad v(y) = u(p+y)\,, \quad y \in \Omega - p\,,$$

on aura

$$v(y) = \int_0^1 G(ty)\cdot y\,\mathrm{d}t$$

d'où

$$\operatorname{grad} v(y) = G(y)\,, \quad y \in \Omega - p\,,$$

ce qui donne

$$\operatorname{grad} u(x) = F(x)\,, \quad x \in \Omega\,. \qquad \blacklozenge$$

REMARQUE. Noter que la définition de u en termes d'une intégrale curviligne est

$$u(x) = \int_{[p\,,x]} F\cdot \mathrm{d}\vec{s}\,.$$

6.1.16 Remarques

Pour un domaine étoilé $\Omega \subset \mathbb{R}^N$, l'énoncé (ii) de la proposition 6.1.15 donne une condition simple, nécessaire et suffisante, pour qu'un champ $F \in C^1(\Omega;\mathbb{R}^N)$ soit conservatif ; si Ω n'est pas étoilé, le problème devient plus compliqué.

Soit $\Omega = \mathbb{R}^2 \setminus \{(0,0)\}$; si

$$\begin{aligned}F(x,y) &= \left(\frac{-y}{x^2+y^2}\,,\,\frac{x}{x^2+y^2}\right)\,, \quad (x,y)\in\Omega \\ &= \big(F_1(x,y)\,,\,F_2(x,y)\big)\end{aligned}$$

on vérifie alors facilement que $D_1F_2 = D_2F_1$ dans Ω ; mais F ne peut pas être conservatif dans Ω car si γ est le lacet défini par un cercle de centre $(0,0)$ et de rayon $r > 0$, parcouru contre le sens des aiguilles d'une montre (analytiquement : $\gamma(t) = r(\cos t, \sin t)$, $0 \leqslant t \leqslant 2\pi$), on aura, après un petit calcul,

$$\int_\gamma F \cdot d\vec{s} = \int_0^{2\pi} dt = 2\pi \neq 0 .$$

Ainsi, Ω est un exemple d'un domaine non étoilé où l'énoncé (ii) de la proposition 6.1.15 est en défaut.

D'autre part, soit $\Omega = \mathbb{R}^N \setminus \{0\}$, $N \geqslant 3$; ici aussi, Ω est un domaine non étoilé mais néanmoins l'énoncé (ii) de la proposition 6.1.15 reste exact dans Ω ; on pourra démontrer ceci directement à l'aide du théorème de Stokes, donné ci-après pour $N = 3$ et dans le chapitre 7 pour N quelconque. En fait, la condition optimale sur le domaine Ω pour la validité de (ii) de la proposition 6.1.15 est celle d'être *simplement connexe* ; cette condition était expliquée intuitivement dans le paragraphe 4.2.9 dans un contexte différent ; elle sera abordée précisément dans le chapitre 7 (voir aussi § 6.2.10).

6.1.17 Courbes

La notion d'une courbe est à la fois une notion très intuitive et très subtile, comme le sont d'ailleurs beaucoup d'autres notions physiques et géométriques. Nous aurions pu éviter entièrement toute mention quelconque des courbes, en nous limitant aux considérations des chemins, trajectoires et variétés différentielles de dimension 1 (définies dans le chapitre 4). Beaucoup d'exposés rigoureux évitent entièrement la définition explicite d'une courbe et un plus grand nombre encore utilise le mot « courbe » sans en donner une définition convaincante. Il est possible de donner une bonne définition topologique du style « une courbe dans $\mathbb{R}^N$ est une partie connexe par arcs qui a la dimension topologique 1 » mais ceci nous entraînerait à une discussion de la dimension topologique sans qu'elle nous soit profitable pour les problèmes analytiques. Néanmoins, tout le monde parle bien d'une courbe parabolique ou d'une courbe exponentielle et il est commode d'avoir une terminologie précise pour couvrir la plupart de ces usages. Nous donnons donc ici une définition d'une courbe qui est utile pour la suite et qui contient la plupart des « courbes usuelles » ; de plus, notre définition ne contredit pas la définition topologique esquissée ci-dessus.

Une ***courbe*** dans $\mathbb{R}^N$ est un ensemble $C \subset \mathbb{R}^N$ tel que $C = \gamma(J)$ où J est un intervalle de $\mathbb{R}$ et où $\gamma : J \to \mathbb{R}^N$ est une trajectoire non constante, C^1 par morceaux ; on dira que C est ***paramétrée*** par γ. Si la courbe C peut être paramétrée par γ qui est simple ou régulière ou régulière par morceaux, on dira que la courbe C est ***simple*** ou ***régulière*** ou ***régulière par morceaux*** et que γ est un ***paramétrage approprié*** de C. Si C est une courbe régulière par morceaux et $x \in C$, on dira que x est un point ***régulier*** de C si dans un paramétrage approprié γ de C, on a $\gamma'(t) \neq 0$ où $x = \gamma(t)$.

On dira que C est une ***courbe de Jordan*** si $C = \gamma(J)$ pour $\gamma : J \to \mathbb{R}^N$ un lacet simple (donc avec J un intervalle compact et γ seulement continue). Une courbe de Jordan est un sous-ensemble compact mais elle n'est pas nécessairement une courbe dans le sens de notre définition ; elle le sera si γ est de plus C^1 par morceaux ; nous parlerons alors d'une courbe de Jordan C^1 par morceaux ; de même, nous parlerons d'une courbe de Jordan régulière par morceaux si le lacet γ est régulier par morceaux.

Si C est une courbe de Jordan C^1 par morceaux, on définira la longueur de C comme la longueur d'un lacet simple γ, C^1 par morceaux, qui possède C comme trace. Plus généralement, si $\gamma : J \to \mathbb{R}^N$ est une trajectoire non constante, C^1 par morceaux, ayant un nombre fini de points multiples et C la courbe donnée par $\gamma(J)$, on définira

$$\text{longueur}(C) = \text{longueur}(\gamma) ;$$

dans ce contexte, on posera

$$\int_C f \, \mathrm{d}s = \int_\gamma f \, \mathrm{d}s , \quad \int_C F \cdot \mathrm{d}\vec{s} = \int_\gamma F \cdot \mathrm{d}\vec{s}$$

où $f : C \to \mathbb{R}$ et $F : C \to \mathbb{R}^N$; il faut être conscient que $\int_C f \, \mathrm{d}s$ ne dépend pas du paramétrage particulier γ utilisé pourvu que celui-ci soit C^1 par morceaux et possède au plus un nombre fini de points multiples ; pour $\int_C F \cdot \mathrm{d}\vec{s}$, le paramétrage choisi impose un sens de parcours ou une ***orientation*** à C et la valeur de $\int_C F \cdot \mathrm{d}\vec{s}$ changera de signe si l'on choisit un autre paramétrage qui renverse l'orientation. Il n'est pas nécessaire d'introduire des définitions plus élaborées sur ce sujet.

On peut démontrer qu'une sous-variété différentielle connexe de dimension 1 de classe C^1 dans $\mathbb{R}^N$ est une courbe dans notre sens ; ce sont les courbes régulières sans points multiples.

6.1.18 Exercices

1 Calculer $\int_\gamma f \, \mathrm{d}s$ dans les exemples suivants (où a, b sont des constantes > 0) :

(i) $f(x,y) = x$; $\gamma(t) = a(\cos t, \sin t)$, $0 \leqslant t \leqslant 2\pi$ (cercle) ;

(ii) $f(x,y) = \sqrt{y}$; $\gamma(t) = a(t - \sin t\,,\, 1 - \cos t)$, $0 \leqslant t \leqslant 2\pi$ (cycloïde) ;

(iii) $f(x,y) = b$; $\gamma(t) = a(\cos^3 t, \sin^3 t)$, $0 \leqslant t \leqslant 2\pi$ (hypocycloïde ou astroïde) ;

(iv) $f(x,y,z) = (x + y + z)$; $\gamma(t) = a(\sin t, \cos t, t)$, $0 \leqslant t \leqslant 2\pi$ (hélice) ;

(v) $f(x,y,z) = e^z$; $\gamma = \big[(1,1,1)\,,\,(-2,3,5)\big]$.

2 Calculer $\int_\gamma F \cdot d\vec{s}$ dans les exemples suivants.

(i) $F(x,y) = (0, xy)$; $\gamma(t) = (\cos t, \sin t)$, $0 \leqslant t \leqslant \pi/2$;
(ii) $F(x,y) = (xy, 0)$; $\gamma(t) = (\cos t, \sin t)$, $0 \leqslant t \leqslant \pi/2$;
(iii) $F(x,y) = (x,y)$; $\gamma(t) = (\cos^3 t, \sin^3 t)$, $0 \leqslant t \leqslant 2\pi$;
(iv) $F(x,y,z) = (y,\ z\cos yz + x,\ y\cos yz)$; $\gamma(t) = (\cos t, \sin t, 1)$, $0 \leqslant t \leqslant 2\pi$;
(v) $F(x,y,z) = (x,y,z)$; $\gamma(t) = (\sin t, \cos t, t)$, $0 \leqslant t \leqslant 2\pi$;
(vi) $F(x,y,z) = (x^2, xy, 1)$; $\gamma(t) = (t, t^2, t^3)$, $0 \leqslant t \leqslant 1$;
(vii) $F(x,y,z) = (1,y,z)$; $\gamma(t) = \big(\cos t, \sin t, |t - \pi/2|\big)$, $0 \leqslant t \leqslant \pi$;
(viii) $F(x,y,z) = (yz, 2y, -x^2)$; $\gamma = [a,b] \vee [b,c] \vee [c,d]$, $a = (0,0,0)$, $b = (0,0,1)$, $c = (0,-3,1)$, $d = (2,-3,1)$;
(ix) $F(x,y,z) = (z/y,\ 0,\ x^2 + y^2 + z^2)$; γ est le chemin de $(0,1,4)$ à $(1,0,6)$ suivant l'intersection du cylindre $x^2 + y^2 = 1$ et du plan $z = 2x + 4$ dans le premier octant ;
(x) $F(x,y,z) = \mu \times x$; $\gamma(t) = \xi \cos t + \eta \sin t$, $0 \leqslant t \leqslant 2\pi$ (μ, ξ, η étant des vecteurs fixes de $\mathbb{R}^3$).

3 Calculer la longueur des courbes suivantes définies dans l'exercice 1 : (i) cercle, (ii) cycloïde, (iii) hypocycloïde, (iv) hélice.

4 Vérifier que les trajectoires suivantes tracent le même cercle sauf que dans (v) la trajectoire ne comprend pas le point $(-1,0)$; calculer les longueurs des trajectoires et expliquer vos résultats.

(i) $\gamma(t) = (\cos t, \sin t)$, $0 \leqslant t \leqslant 2\pi$.
(ii) $\gamma(t) = (\sin t, \cos t)$, $0 \leqslant t \leqslant 2\pi$;
(iii) $\gamma(t) = (\cos nt, \sin nt)$, $0 \leqslant t \leqslant 2\pi$ (n entier $\neq 0$) ;
(iv) $\gamma(t) = (\cos 2\pi t^2,\ \sin 2\pi t^2)$, $-1 \leqslant t \leqslant 1$;
(v) $\gamma(t) = \left(\dfrac{1-t^2}{1+t^2},\ \dfrac{2t}{1+t^2}\right)$, $-\infty < t < \infty$;
(vi) $\gamma(t) = \begin{cases} (t, \sqrt{1-t^2}) & \text{si } -1 \leqslant t \leqslant 1 \\ \big(2-t, -\sqrt{1-(2-t)^2}\big) & \text{si } 1 \leqslant t \leqslant 3. \end{cases}$

5 Une courbe dans $\mathbb{R}^2$ donnée par une équation en coordonnées polaires $r = \rho(\theta)$, $\theta \in J$, J étant un intervalle de $\mathbb{R}$ et $\rho : J \to [0, \infty[$ continue, peut être paramétrée par $\gamma(t) = \big(\rho(t)\cos t,\ \rho(t)\sin t\big)$, $t \in J$; si ρ est C^1 par morceaux, démontrer que la longueur de γ est donnée par

$$\int_J \sqrt{\rho^2(t) + \rho'^2(t)}\ dt.$$

En déduire la longueur de la cardioïde d'équation polaire $r = a(1 + \cos\theta)$, $0 \leqslant \theta \leqslant 2\pi$.

6 Vérifier en passant aux coordonnées cartésiennes de $\mathbb{R}^2$ que l'équation polaire $r = k\varepsilon/(1 - \varepsilon\cos\theta)$ $(k > 0,\ \varepsilon > 0)$ correspond à

(i) une parabole si $\varepsilon = 1$,
(ii) une ellipse si $0 < \varepsilon < 1$,
(iii) une hyperbole si $\varepsilon > 1$;

le cas limite de $\varepsilon \to 0$ et $k\varepsilon \to a > 0$ donne un cercle.

7 La longueur d'une ellipse paramétrée par $\gamma(t) = (a\cos t, b\sin t)$, $0 \leqslant t \leqslant 2\pi$ $(a > b)$, est donnée par $4aE(\varepsilon; \pi/2)$ où

$$E(\varepsilon;\pi/2) = \int_0^{\pi/2} (1 - \varepsilon^2 \sin^2 t)^{1/2}\, \mathrm{d}t\,, \quad \varepsilon = \left(1 - \frac{b^2}{a^2}\right)^{1/2}$$

$E(\varepsilon;\pi/2)$ est un exemple d'une intégrale elliptique.

8 Vérifier que les deux trajectoires dans $\mathbb{R}^2$ données par $t \mapsto (a\sec t, b\,\mathrm{tg}\, t)$, $a > 0$, $b > 0$, lorsque t est dans $]-\pi/2, \pi/2[$ ou dans $]\pi/2, 3\pi/2[$, donnent les deux branches de l'hyperbole d'équation cartésienne $x^2/a^2 - y^2/b^2 = 1$. Vérifier que la trajectoire donnée par $t \mapsto (a\,\mathrm{ch}\, t, b\,\mathrm{sh}\, t)$, $t \in \mathbb{R}$, trace la branche de l'hyperbole ci-dessus où $x > 0$.

9 La lemniscate de Bernoulli est une courbe dans $\mathbb{R}^2$ donnée par l'équation cartésienne $(x^2 + y^2)^2 = 2a^2(x^2 - y^2)$, $a > 0$; vérifier que son équation en coordonnées polaires est $r^2 = 2a^2\cos 2\theta$ où θ varie dans $[-\pi/4, \pi/4] \cup [3\pi/4, 5\pi/4]$. En déduire que la lemniscate est paramétrée par $\gamma_1 \vee \gamma_2$ où

$$\gamma_1 = \gamma|_{[-\pi/4,\,\pi/4]} \text{ et } \gamma_2 = \gamma|_{[3\pi/4,\,5\pi/4]} \quad \text{avec} \gamma(t) = \big(\rho(t)\cos t\,,\ \rho(t)\sin t\big)$$

où $\rho(t) = a\sqrt{2\cos 2t}$; chaque γ_i est un lacet, γ_1 paramètre la boucle de la lemniscate où $x \geqslant 0$ et γ_2 celle où $x \leqslant 0$; les lacets γ_1, γ_2 ne sont ni réguliers par morceaux ni C^1 par morceaux car $\|\gamma_i'(t)\|$ n'est pas bornée ; chaque γ_i est rectifiable, la longueur de chaque γ_i étant égale à

$$2a\sqrt{2}\int_0^{\pi/4} \frac{\mathrm{d}t}{\sqrt{\cos 2t}}$$

qui est un autre exemple d'une intégrale elliptique. Un paramétrage de classe C^∞ de la lemniscate est donné par

$$\xi(u) = a\sqrt{2}\left(\frac{u(1+u^2)}{1+u^4}\,,\ \frac{u(1-u^2)}{1+u^4}\right), \quad -\infty < u < \infty\,;$$

noter que

$$\xi(0,0) = (0,0)\,, \quad \lim_{u\to\infty} \xi(u) = (0,0)\,, \quad \lim_{u\to-\infty} \xi(u) = (0,0)$$

et que ξ donne un paramétrage de la lemniscate sans point multiple (alors que $(0,0)$ est « clairement » un point multiple pour la courbe) ; lorsque u varie de $-\infty$ à ∞, $\xi(u)$ trace la lemniscate de telle manière que chaque point de la courbe correspond à une unique valeur de $u \in \mathbb{R}$. Vérifier que $\xi : \mathbb{R} \to \mathbb{R}^2$ est telle que $\xi'(u) \neq (0,0)$ pour tout u et que ξ est une application injective continue qui n'est pas un homéomorphisme entre $\mathbb{R}$ et $\xi(\mathbb{R}^2)$. Posons

$$\delta = \xi|_{[-1,1]} \quad \text{et} \quad \eta(u) = a\sqrt{2}\left(\frac{-u(1+u^2)}{1+u^4}, \frac{u(1-u^2)}{1+u^4}\right), \quad -1 \leqslant u \leqslant 1;$$

vérifier que $\delta \vee \eta$ donne un lacet de classe C^1 qui trace la lemniscate avec $(0,0)$ comme unique point multiple (double). Calculer $\delta'(-1)$, $\delta'(0)$, $\delta'(1)$, $\eta'(-1)$, $\eta'(0)$, $\eta'(1)$ pour vérifier le bon sens de parcours.

10 Soit $C = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$; si $F : C \to \mathbb{R}^2$ est un champ vectoriel continu et $\gamma_n(t) = (\cos nt, \sin nt)$, $0 \leqslant t \leqslant 2\pi$, n étant un entier $\neq 0$, vérifier que

$$\int_{\gamma_n} F \cdot d\vec{s} = n \int_{\gamma_1} F \cdot d\vec{s}.$$

11 Si $f : \mathbb{R} \to \mathbb{R}$ est continue, voir directement que

$$\int_\gamma f(x)\,dx = \int_0^{2\pi} f(x)\,dx$$

où $\gamma(t) = (t\cos t, t\sin t, t)$, $0 \leqslant t \leqslant 2\pi$; voir que le champ vectoriel $F(x,y,z) = (f(x),0,0)$ provient d'un potentiel u donné par $u(x,y,z) = g(x)$ où g est une primitive de f d'où suit aussi l'égalité des deux intégrales.

12 Soit $F \in C(\Omega;\mathbb{R}^N)$, Ω un domaine dans $\mathbb{R}^N$; supposons que

$$\int_\gamma F \cdot d\vec{s} = f(y) - f(x)$$

pour tout chemin polygonal simple γ reliant x à y dans Ω où $f : \Omega \to \mathbb{R}$. Démontrer que $f \in C^1(\Omega)$ et que $F = \nabla f$ dans Ω.

13 Soient F, G deux champs vectoriels continus dans un domaine Ω de $\mathbb{R}^N$; si $\int_\gamma F \cdot d\vec{s} = \int_\gamma G \cdot d\vec{s}$ pour tout chemin polygonal simple dans Ω, démontrer que $F = G$.

14 Soit $f :]0,\infty[\to \mathbb{R}$ une fonction continue. Vérifier que le champ vectoriel (« force centrale ») F donné par $F(x) = f(\|x\|)x$, $x \in \mathbb{R}^N \setminus \{0\}$, provient d'un potentiel $u(x) = g(\|x\|)$ où g est une primitive de la fonction $r \mapsto rf(r)$, $r > 0$.

15 Parmi les champs vectoriels suivants déterminer ceux qui proviennent d'un potentiel et dans ces derniers cas donner un potentiel :

(i) $F(x,y) = (x^2y, \sin xy)$, $(x,y) \in \mathbb{R}^2$;
(ii) $F(x,y) = (2xy, x^2 + 3y^2)$, $(x,y) \in \mathbb{R}^2$;
(iii) $F(x,y,z) = (y,z,x)$, $(x,y,z) \in \mathbb{R}^3$;
(iv) $F(x,y,z) = (x,y,z)$, $(x,y,z) \in \mathbb{R}^3$;
(v) $F(x) = a$, $x \in \mathbb{R}^N$ (a vecteur constant dans $\mathbb{R}^N$) ;
(vi) $F(x) = \big(f_1(x_1), \ldots, f_N(x_N)\big)$, $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$, $f_i \in C(\mathbb{R})$, $i = 1, \ldots, N$.

16 Calculer la masse et le centre de gravité d'un fil métallique de densité constante ρ d'équation $x = \cos t$, $y = \sin t$, $z = ht$, $0 \leqslant t \leqslant T$ $(h > 0)$.

17 Calculer :

(i) $\displaystyle\int_\gamma \{x^n y^{n+1}\,\mathrm{d}x + x^{n+1}y^n\,\mathrm{d}y\}$
où γ est l'arc de la parabole d'équation $y^2 = x$ de $(1,-1)$ à $(1,1)$ et où n est un entier $\geqslant 0$;

(ii) $\displaystyle\int_\gamma \left\{\frac{x}{x^2+y^2}\,\mathrm{d}x + \frac{y}{x^2+y^2}\,\mathrm{d}y\right\}$ où γ est le segment de $(1,0)$ à $(1,1)$;

(iii) $\displaystyle\int_\gamma \{f(x)\,\mathrm{d}x + g(x,y,z)\,\mathrm{d}y - g(x,y,z)\,\mathrm{d}z\}$
où $\gamma(t) = (\cos t, \sin t, \sin t)$, $0 \leqslant t \leqslant 2\pi$ et où $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R}^3 \to \mathbb{R}^3$ sont continues ;

(iv) $\displaystyle\int_\gamma \{(x+y)\,\mathrm{d}x + (x-y)\,\mathrm{d}y\}$
où γ est l'arc de la parabole d'équation $y = x^2$ de $(-1,1)$ à $(1,1)$;

(v) $\displaystyle\int_\gamma xy^2\,\mathrm{d}y$
où γ est un paramétrage de l'ellipse d'équation $4x^2 + y^2 = 4$ de sens contraire au sens des aiguilles d'une montre.

18 Soit $T : \mathbb{R}^N \to \mathbb{R}^N$ une application linéaire ; démontrer que $F(x) = Tx$, $x \in \mathbb{R}^N$, est un champ vectoriel conservatif si et seulement si T est symétrique et déterminer un potentiel pour F dans ce cas. En déduire que si $F(x) = a \times x$, $x \in \mathbb{R}^3$, alors F est conservatif si et seulement si $a = 0$.

19 Soit $C = \{(0,y) : 0 \leqslant y < \infty\} \cup \{(x,0) : 0 \leqslant x < \infty\} \subset \mathbb{R}^2$. Vérifier que C est paramétré par $\gamma(t) = \big(\gamma_1(t), \gamma_2(t)\big)$, $t \in \mathbb{R}$ de trois façons différentes selon les définitions suivantes :

(i) $\gamma_1(t) = \max(t,0)$, $\gamma_2(t) = \max(-t,0)$;
(ii) $\gamma_1(t) = t^2\,1_{[0,\infty[}(t)$, $\gamma_2(t) = t^2\,1_{]-\infty,0]}(t)$;

(iii) $\gamma_1(t) = \varphi(t)\,1_{[0,\infty[}(t)$, $\gamma_2(t) = \varphi(t)\,1_{]-\infty,0]}(t)$
où $\varphi(t) = \mathrm{e}^{-t^{-2}}$ si $t \neq 0$, $\varphi(0) = 0$.

Vérifier que dans (i) $\gamma'(t)$ existe pour tout $t \neq 0$, γ étant C^1 par morceaux, dans (ii) γ est de classe C^1 et dans (iii) γ est de classe C^∞. Dans (ii) et (iii) $\gamma'(0) = (0,0)$; conclure (ce qui est géométriquement évident) que C est une courbe, régulière partout sauf en $(0,0)$, sans point multiple.

REMARQUE. On peut montrer que toute courbe dans notre sens peut être paramétrée de manière C^∞.

6.2 THÉORÈME DE GREEN

Dans cette section, nous donnons une formule (6.14) qui relie une intégrale curviligne d'un champ vectoriel dans $\mathbb{R}^2$ à une intégrale associée, étendue sur un domaine de $\mathbb{R}^2$. La formule s'appelle aussi quelquefois la ***formule de Green-Riemann*** où le ***théorème de Green***.

6.2.1 Explications intuitives

Soit Ω un domaine borné de $\mathbb{R}^2$ dont la frontière est une courbe de Jordan C, régulière par morceaux ; soit $\gamma : [a, b] \to \mathbb{R}^2$ un paramétrage de C tel que γ est un lacet simple régulier par morceaux avec $\gamma([a, b]) = C$.

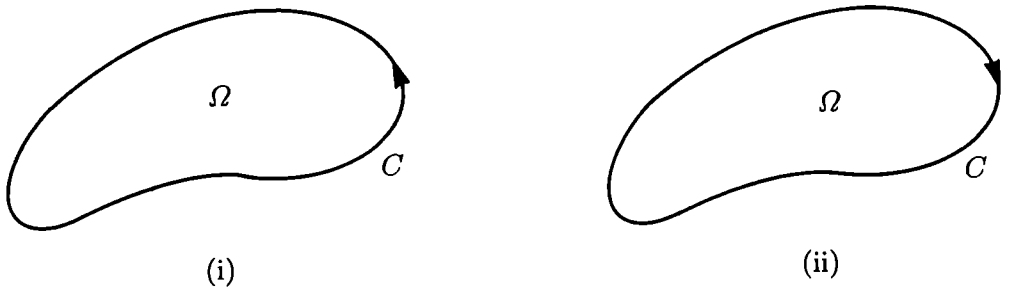


Fig. 6.2

Si le sens de parcours de γ est comme dans (i) de la figure 6.2, on dira que C est ***orientée positivement*** par γ, si le sens de parcours de γ est comme dans (ii) de la figure 6.2, on dira que C est ***orientée négativement*** par γ. Dans le premier cas, on dit aussi que l'on parcourt C de telle manière que Ω reste toujours à la gauche.

Supposons que γ oriente C positivement ; si P, Q sont deux fonctions réelles dans $C^1(\overline{\Omega})$, on a la formule suivante (*théorème de Green*) :

$$\int_\gamma (P\,\mathrm{d}x + Q\,\mathrm{d}y) = \iint_\Omega \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y\,. \tag{6.14}$$

La formule (6.14) est valable dans des situations plus générales ; on peut l'appliquer aux domaines bornés $\Omega \subset \mathbb{R}^2$ tels que $\operatorname{Fr}\Omega$ est la réunion d'un nombre fini de courbes de Jordan $C_1, \ldots, C_n$, deux à deux disjointes, régulières par morceaux et orientées, comme dans la figure 6.3.

Fig. 6.3

Si $\gamma_1, \ldots, \gamma_n$ sont les lacets simples, réguliers par morceaux qui paramétrisent $C_1, \ldots, C_n$, on écrira $\gamma = \gamma_1 \vee \cdots \vee \gamma_n$ pour leur réunion et on pose

$$\int_\gamma F \cdot \mathrm{d}\vec{s} = \int_\gamma (P\,\mathrm{d}x + Q\,\mathrm{d}y) = \sum_{i=1}^{n} \int_{\gamma_i} (P\,\mathrm{d}x + Q\,\mathrm{d}y)$$

où $F = (P, Q)$; avec cette notation, (6.14) reste valable.

Notons que l'orientation de la courbe extérieure C_1 de la figure 6.3 est dans le sens contraire du sens des aiguilles d'une montre alors que celle des autres C_2, C_3, ... est dans l'autre sens (sens des aiguilles d'une montre).

La formulation mathématique de ce qui est dit ci-dessus est un peu longue ; nous l'établirons dans la suite de cette section. Sa démonstration est contenue dans celle d'un théorème plus général (théorème de Stokes général) donné dans le chapitre 7.

Notons ici comme un aide-mémoire le calcul formel suivant :

(i) $\dfrac{\partial Q}{\partial x} - \dfrac{\partial P}{\partial y} = \det \begin{bmatrix} \partial/\partial x & \partial/\partial y \\ P & Q \end{bmatrix}$

(ii) si $F(x, y, z) = \big(P(x, y),\, Q(x, y),\, 0\big)$,

$$\operatorname{rot} F = \left(0,\, 0,\, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right).$$

6.2.2 Une vérification

Dans ce paragraphe nous vérifions le théorème de Green (6.14) pour une classe de domaines bornés Ω dans $\mathbb{R}^2$; les points de $\mathbb{R}^2$ sont notés (x, y).

Nous dirons qu'un domaine Ω est ***x-simple*** si, pour chaque $x \in \mathbb{R}$, la x-section de Ω est ou bien vide ou bien un intervalle ouvert ; on définit un domaine ***y-simple*** de manière analogue en utilisant les y-sections de Ω. Il est clair que si Ω est un domaine borné, x-simple, alors Ω est de la forme suivante :

$$\Omega = \big\{(x, y) : a < x < b,\ f(x) < y < g(x)\big\}$$

avec a, b, $f(x)$, $g(x)$ dans $\mathbb{R}$; dans ce cas, en supposant f, g continues dans $[a, b]$, on aura

$$\operatorname{Fr}\Omega = F_1 \cup F_2 \cup F_3 \cup F_4$$

où

$$F_1 = \{(x, f(x)) : a \leqslant x \leqslant b\}\ , \quad F_2 = \{(b, y) : f(b) \leqslant y \leqslant g(b)\}$$
$$F_3 = \{(x, g(x)) : a \leqslant x \leqslant b\}\ , \quad F_4 = \{(a, y) : f(a) \leqslant y \leqslant g(a)\}\ .$$

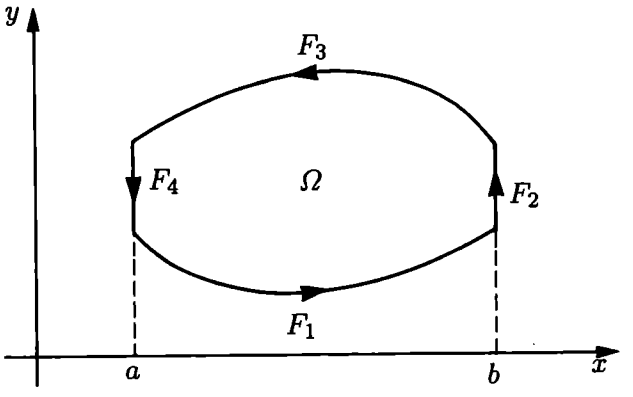


Fig. 6.4

On convient que $F_2 = \emptyset$ si $f(b) = g(b)$ et $F_4 = \emptyset$ si $f(a) = g(a)$. Les fonctions f, g peuvent être arbitraires sauf que $f(x) < g(x)$ si $a < x < b$ et elles sont continues dans $[a, b]$; on pourra paramétrer $\operatorname{Fr}\Omega$ comme un lacet

$$\gamma = \gamma_1 \vee \gamma_2 \vee \gamma_3 \vee \gamma_4$$

où

$$\gamma_1(x) = (x, f(x))\ ,\ a \leqslant x \leqslant b\ ; \quad \gamma_2(y) = (b, y)\ ,\ f(b) \leqslant y \leqslant g(b)$$
$$\widetilde{\gamma}_3(x) = (x, g(x))\ ,\ a \leqslant x \leqslant b\ ; \quad \widetilde{\gamma}_4(y) = (a, y)\ ,\ f(a) \leqslant y \leqslant g(a)\ ,$$

$\widetilde{\gamma}_3$ étant opposé à γ_2, $\widetilde{\gamma}_4$ opposé à γ_4. Le lacet γ sera C^1 par morceaux si les fonctions f et g le sont ; en fait, pour calculer $\int_\gamma P\,\mathrm{d}x$ on n'aura pas besoin de cette condition. Si $P \in C^1(\overline{\Omega})$, on aura

$$\begin{aligned}\int_\gamma P\,\mathrm{d}x &= \int_{\gamma_1} P\,\mathrm{d}x + \int_{\gamma_2} P\,\mathrm{d}x + \int_{\gamma_3} P\,\mathrm{d}x + \int_{\gamma_4} P\,\mathrm{d}x \\ &= \int_a^b P\big(x, f(x)\big)\,\mathrm{d}x - \int_a^b P\big(x, g(x)\big)\,\mathrm{d}x \\ &= \int_a^b \left\{ -\int_{f(x)}^{g(x)} \frac{\partial P}{\partial y}\,\mathrm{d}y \right\} \mathrm{d}x \\ &= -\iint_\Omega \frac{\partial P}{\partial y}\,\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

en remarquant que les intégrales sur γ_2 et γ_4 sont nulles et que l'intégrale sur γ_3 est l'opposé de l'intégrale sur $\widetilde{\gamma}_3$. Ceci démontre (6.14) dans le cas $Q = 0$ et un domaine borné Ω x-simple. Un raisonnement analogue établit (6.14) dans le cas $P = 0$ et pour un domaine borné Ω y-simple. Si le domaine borné Ω est à la fois x-simple et y-simple, on aura alors (6.14) pour tout P, Q dans $C^1(\Omega)$.

Beaucoup de domaines bornés dans $\mathbb{R}^2$ peuvent être représentés comme une réunion finie des domaines à la fois x-simples et y-simples et on peut établir (6.14) dans ces cas à partir de cette propriété.

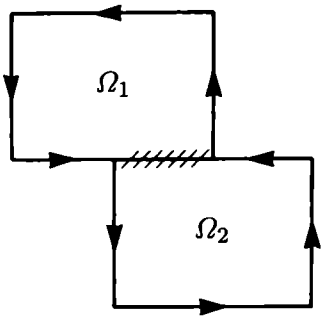


Fig. 6.5

Par exemple si $\Omega = \Omega_1 \cup \Omega_2$ (fig. 6.5) la validité de (6.14) pour Ω_1 et Ω_2 entraîne la validité de cette formule pour Ω car l'intégrale sur le chemin hachuré est annulée par son parcours deux fois dans les directions opposées. Nous ne développons pas ce raisonement plus loin car il devient fastidieux à écrire en détail ; une démonstration rigoureuse d'une formule plus générale apparaît dans le chapitre 7.

Pour énoncer le théorème de Green dans une forme précise, nous devons introduire plusieurs considérations que nous abordons maintenant.

6.2.3 Théorème de Jordan

Ce théorème difficile n'est pas indispensable pour nous, mais il permet de clarifier certaines idées géométriques, intuitivement évidentes mais subtiles à préciser ; aussi, il nous donne une terminologie mathématique commode.

Soit C une courbe de Jordan dans $\mathbb{R}^2$; rappelons que ceci veut dire que $C = \gamma(J)$ où $\gamma : J \to \mathbb{R}^2$ est un lacet simple, J étant un intervalle compact de $\mathbb{R}$. *Le théorème de Jordan dit que*

$$\mathbb{R}^2 = \Omega \cup C \cup \Omega'$$

où Ω est un domaine borné, Ω' est un domaine non borné et Ω, C, Ω' sont deux à deux disjoints ; les domaines Ω, Ω' sont uniquement déterminés par C et

$$C = \operatorname{Fr} \Omega = \operatorname{Fr} \Omega' .$$

Nous dirons que Ω est le ***domaine intérieur*** défini par C et que Ω' est le ***domaine extérieur*** défini par C. Notons que $\Omega' = \operatorname{ext} \Omega = \left(\overline{\Omega}\right)^c$ et $\overline{\Omega} = \Omega \cup C$.

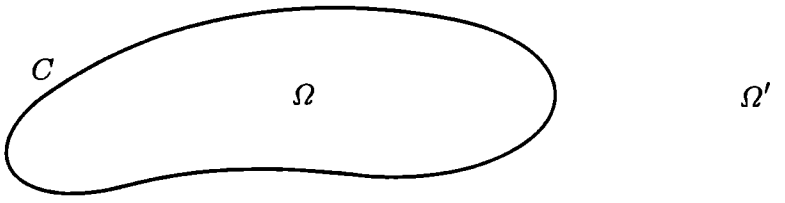


Fig. 6.6

Pour les courbes de Jordan qui ne sont pas trop compliquées (fig. 6.6), la validité du théorème de Jordan est visuellement évidente. En fait, si le lacet γ qui définit C est de classe C^1 et régulier, la démonstration du théorème n'est pas trop difficile ; mais la démonstration dans le cas général est longue si l'on veut procéder sans l'aide des outils de la topologie algébrique, un sujet que nous n'abordons pas dans ce livre. Une preuve élémentaire utilisant les fonctions complexes est donnée dans Dieudonné [13, tome 1, p. 261] ; une démonstration utilisant un minimum de topologie algébrique est donnée dans Newman [31] qui établit aussi une version appropriée du théorème de Jordan dans $\mathbb{R}^N$, $N \geqslant 2$, dû à Brouwer.

Un critère pratique pour déterminer si un point p dans le complémentaire de C est dans Ω ou dans Ω', est le suivant : s'il existe une demi-droite d'origine p qui intersecte C en un nombre fini impair de points, alors $p \in \Omega$; s'il existe une demi-droite d'origine p qui intersecte C en un nombre fini pair (éventuellement 0) de points, alors $p \in \Omega'$. On pourra vérifier l'utilité de ce critère dans un diagramme de C plus compliqué que celui de la figure 6.6.

6.2.4 Domaines réguliers dans $\mathbb{R}^2$ et orientation de leurs frontières

Soit $\gamma :]a, b[\to \mathbb{R}^2$ une trajectoire simple régulière et C la courbe simple régulière tracée par $\gamma = (\gamma_1, \gamma_2)$; si $x \in C$ et $\gamma(t) = x$, nous dirons que l'***espace tangent*** de C au point x est le sous-espace vectoriel de $\mathbb{R}^2$ de dimension 1 engendré par le vecteur non nul $\gamma'(t)$ et nous le noterons par T_x ; un vecteur ***normal*** à C au point x est un vecteur $\xi \in \mathbb{R}^2$ tel que $\xi \cdot \gamma'(t) = 0$. Ainsi, $\xi = \big(\gamma_2'(t), -\gamma_1'(t)\big)$ est un vecteur normal à C au point x et tous les autres vecteurs normaux à C au point x sont les multiples de ξ ; parmi ceux-ci, les deux vecteurs normaux unitaires seront notés par $\pm\nu(x)$ de sorte que $\pm\nu(x) = \pm\xi/\|\gamma'(t)\|$. Il est clair qu'un reparamétrage de C ne change ni T_x ni les vecteurs normaux $\pm\nu(x)$ car si $\theta :]c, d[\to]a, b[$ est un C^1-difféomorphisme et $x = \gamma(t) = \gamma\big(\theta(u)\big)$, on a

$$\gamma'\big(\theta(u)\big)\theta'(u) = \gamma'(t)$$

et, puisque $\theta'(u) \neq 0$, les espaces engendrés par les vecteurs $\gamma'(t)$ et $\gamma'\big(\theta(u)\big)$ sont les mêmes ; θ peut éventuellement renverser la direction de $\gamma'(t)$ mais ceci ne nous concerne pas ici. La discussion précédente est aussi applicable aux

points *réguliers* d'une courbe de Jordan C régulière par morceaux ; rappelons que si C est paramétrée par $\gamma : [a, b] \to \mathbb{R}^2$, un chemin simple fermé C^1 par morceaux avec $\gamma([a, b]) = C$ et si $x = \gamma(t)$, $a < t < b$, est tel que $\gamma'(t)$ existe et $\gamma'(t) \neq 0$ alors x s'appelle un point ***régulier*** de C ; si $x = \gamma(a) = \gamma(b)$, x est un point régulier si $\gamma'(a)$, $\gamma'(b)$ existent (comme dérivées unilatérales) et $\gamma'(a) = \gamma'(b) \neq 0$; on vérifie facilement que la notion d'un point régulier est indépendante du paramétrage γ choisi. Notons qu'une courbe régulière par morceaux possède au plus un nombre fini de points non réguliers.

Soit Ω un domaine borné dans $\mathbb{R}^2$ dont la frontière est une courbe de Jordan C, régulière par morceaux ; nous dirons que C est orientée ***positivement*** par un paramétrage $\gamma : [a, b] \to \mathbb{R}^2$ si à chaque point régulier x de C avec $x = \gamma(t) = (\gamma_1(t), \gamma_2(t))$, le vecteur normal unitaire $\nu(x)$ donné par la normalisation de $(\gamma_2'(t), -\gamma_1'(t))$ possède la propriété suivante : il existe un $\varepsilon_0 > 0$ tel que

$$x + \varepsilon\nu(x) \in \operatorname{ext}\Omega, \quad x - \varepsilon\nu(x) \in \Omega, \quad 0 < \varepsilon < \varepsilon_0$$

où $\operatorname{ext}\Omega = (\overline{\Omega})^c$; on dira que $\nu(x)$ est la normale unitaire (au point x de C) ***dirigée vers l'extérieur*** (de Ω) et $-\nu(x)$ la normale unitaire ***dirigée vers l'intérieur***. Notons que les deux vecteurs $\{\nu(x), \gamma'(t)\}$ forment une base positive de $\mathbb{R}^2$ car

$$\det \begin{bmatrix} \gamma_2'(t) & \gamma_1'(t) \\ -\gamma_1'(t) & \gamma_2'(t) \end{bmatrix} = \|\gamma'(t)\|^2 > 0.$$

L'orientation positive de C définie ci-dessus est la formulation analytique de l'idée géométrique d'un parcours de C de telle manière que le domaine intérieur Ω soit toujours « à notre gauche ».

Si Ω est le disque unité dans $\mathbb{R}^2$ donné par

$$\{(x_1, x_2) : x_1^2 + x_2^2 < 1\}$$

et si l'on paramétrise $\operatorname{Fr}\Omega = C$ par

$$\gamma(t) = (\cos t, \sin t), \quad 0 \leqslant t \leqslant 2\pi,$$

on voit facilement que le cercle C est orienté positivement par γ selon la définition proposée à la figure 6.7 ; cette orientation est souvent décrite comme étant « de sens contraire au mouvement des aiguilles d'une montre ».

On peut démontrer qu'il est toujours possible d'orienter $\operatorname{Fr}\Omega = C$ (Ω domaine borné, C une courbe de Jordan régulière par morceaux) positivement par un choix approprié du paramétrage de C ; ceci paraît intuitivement évident mais la démonstration n'est pas banale ; elle sera discutée dans un cadre plus général dans le chapitre 7. Si γ est un paramétrage positif de C, le chemin opposé $\tilde{\gamma}$ donne une orientation de C que nous appellerons ***négative*** ; $\tilde{\gamma}$ parcourt C avec le domaine Ω « à notre droite » et « dans le sens du mouvement des aiguilles d'une montre ».

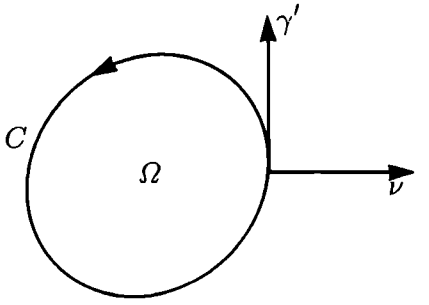


Fig. 6.7

Un cas ***simple*** d'un domaine ***régulier*** dans $\mathbb{R}^2$ est un domaine $\Omega \subset \mathbb{R}^2$ tel que $\operatorname{Fr}\Omega = C$ est une courbe de Jordan régulière par morceaux ; nous avons déjà expliqué ci-dessus la notion d'orientation positive ou négative de $\operatorname{Fr}\Omega$ d'un tel domaine. Le cas général d'un domaine ***régulier*** dans $\mathbb{R}^2$ est un domaine borné $\Omega \subset \mathbb{R}^2$ de la forme suivante :

$$\Omega = D_1 \setminus (\overline{D}_2 \cup \cdots \cup \overline{D}_n)$$

où chaque D_i est un domaine régulier simple avec $C_i = \operatorname{Fr} D_i$ une courbe de Jordan régulière par morceaux et $\overline{D}_2, \ldots, \overline{D}_n$ sont deux à deux disjoints et contenus dans D_1. Il n'est pas difficile de voir que

$$\operatorname{Fr}\Omega = C_1 \cup C_2 \cup \cdots \cup C_n .$$

Supposons que γ_i est un paramétrage de C_i, $i = 1, \ldots, n$; nous dirons que $\gamma = \gamma_1 \vee \cdots \vee \gamma_n$ oriente $\operatorname{Fr}\Omega$ ***positivement*** si γ_1 oriente C_1 positivement et si γ_i oriente C_i négativement pour $i = 2, \ldots, n$; si γ_1 oriente C_1 négativement et si γ_i oriente C_i positivement pour $i = 2, \ldots, n$, nous dirons que γ oriente $\operatorname{Fr}\Omega$ ***négativement***. Si F est un champ vectoriel défini sur $\operatorname{Fr}\Omega$, nous écrivons

$$\int_{\operatorname{Fr}\Omega} F \cdot \mathrm{d}\vec{s} = \sum_{i=1}^{n} \int_{\gamma} F \cdot \mathrm{d}\vec{s}$$

après avoir précisé l'orientation choisie pour $\operatorname{Fr}\Omega$. Si $F = (P, Q)$, nous écrivons aussi

$$\int_{\operatorname{Fr}\Omega} (P\,\mathrm{d}x + Q\,\mathrm{d}y)$$

au lieu de $\int_{\operatorname{Fr}\Omega} F \cdot \mathrm{d}\vec{s}$.

6.2.5 Proposition (*formulation précise du théorème de Green*)

Soit Ω un domaine régulier dans $\mathbb{R}^2$. Si P, Q sont deux fonctions réelles dans $C^1(\overline{\Omega})$, alors

$$\int_{\operatorname{Fr}\Omega} (P\,\mathrm{d}x + Q\,\mathrm{d}y) = \iint_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y$$

où $\operatorname{Fr}\Omega$ est orientée positivement.

REMARQUES: Une version plus générale de cette proposition sera démontrée dans le chapitre 7 comme cas particulier du théorème de Stokes général dans $\mathbb{R}^N$. La démonstration dans le cas d'un domaine Ω qui est x-y-simple, donnée dans le paragraphe 6.2.2, illustre déjà l'importance de la condition d'orientation positive de $\operatorname{Fr}\Omega$. Il faut ajouter que la terminologie *domaine régulier* n'est pas universelle ; il faut comprendre que pour la validité du théorème de Green, il faut non seulement imposer certaines conditions sur le champ vectoriel $F = (P, Q)$ mais aussi sur Ω et $\operatorname{Fr}\Omega$ (§ 6.1.16 et 6.2.9). Il y a des contre-exemples plus subtiles où le domaine Ω n'est pas régulier dans notre sens ; cf., Whitney [43, p. 100] ; Whitney introduit la notion du « domaine standard » (de $\mathbb{R}^n$) dans ce contexte ; c'est une notion plus générale que celle du domaine régulier. Dans l'exposé de Cartan [8, p. 67], on parle du « compact à bord » pour les ensembles $\overline{\Omega}$ où Ω est un domaine régulier dans notre sens. Les conditions les plus générales pour la validité du théorème de Green et ses analogues sont données dans le livre de Federer [18].

Nous donnons maintenant quelques applications du théorème de Green.

6.2.6 Calcul d'aire

En prenant $P(x, y) = -y$, $Q(x, y) = 0$ dans le théorème de Green, on obtient

$$\lambda_2(\Omega) = -\int_\gamma y\,\mathrm{d}x\ ;$$

de même, avec $P(x, y) = 0$, $Q(x, y) = x$, on obtient

$$\lambda_2(\Omega) = \int_\gamma x\,\mathrm{d}y$$

d'où

$$\lambda_2(\Omega) = \int_\gamma x\,\mathrm{d}y = -\int_\gamma y\,\mathrm{d}x = \frac{1}{2}\int_\gamma (x\,\mathrm{d}y - y\,\mathrm{d}x)\,. \tag{6.15}$$

On écrit aussi (6.15) comme suit :

$$\lambda_2(\Omega) = \frac{1}{2}\int_C (x\,\mathrm{d}y - y\,\mathrm{d}x)\,,$$

avec pour convention que $C = \operatorname{Fr}\Omega$ est orientée positivement par rapport à Ω.

Plusieurs exercices dans le paragraphe 6.2.11 illustrent l'utilité de (6.15) dans le calcul des aires dans $\mathbb{R}^2$.

6.2.7 Formule de changement de variables dans $\mathbb{R}^2$

Soit $\varphi : U \to V$ un difféomorphisme de classe C^2 entre les domaines U, V de $\mathbb{R}^2$. Il est alors clair que $\det \varphi'$ ne change pas de signe dans U ; dans ce paragraphe, nous obtenons, comme une application du théorème de Green, non

seulement la formule pour le changement d'aire dans $\mathbb{R}^2$ sous l'application φ mais aussi que φ préserve les orientations si $\det \varphi' > 0$ et que φ renverse les orientations si $\det \varphi' < 0$.

Supposons que D soit un domaine borné contenu dans U avec $\operatorname{Fr} D = C$, où C est une courbe de Jordan régulière par morceaux ; soit $\varphi(D) = E$; alors E est un domaine borné dans V et $\operatorname{Fr} E = F$ où $F = \varphi(C)$; F sera aussi une courbe de Jordan régulière par morceaux. Nous orientons C positivement par rapport à D et posons

$$\varphi(u,v) = \big(a(u,v)\,,\, b(u,v)\big)\,, \quad (u,v) \in U\,.$$

Alors

$$\int_F x\,\mathrm{d}y = \int_C a\{b'_u\,\mathrm{d}u + b'_v\,\mathrm{d}v\}$$

où F est orienté suivant $\varphi(C)$ en ce sens que si $t \mapsto \big(\alpha(t), \beta(t)\big)$ est un paramétrage de C, celui de F sera

$$t \mapsto \big(a\big(\alpha(t),\beta(t)\big)\,,\, b\big(\alpha(t),\beta(t)\big)\big) = \varphi\big(\alpha(t),\beta(t)\big)$$

d'où un petit calcul établit l'égalité entre les deux intégrales. Selon le théorème de Green appliqué sur la seconde intégrale, on obtient, après un calcul direct (utilisant l'égalité des dérivées mixtes de b qui est de classe C^2), que

$$\int_C a\{b'_u\,\mathrm{d}u + b'_v\,\mathrm{d}v\} = \iint_D \frac{\partial(a,b)}{\partial(u,v)}\,\mathrm{d}u\,\mathrm{d}v\,.$$

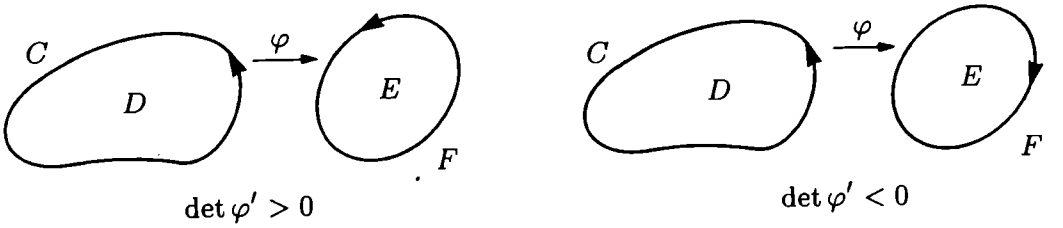
Puisque (selon § 6.2.6)

$$\int_F x\,\mathrm{d}y = \pm\lambda_2(E)\,,$$

avec le signe + si l'orientation induite de F est positive et avec le signe − si l'orientation induite de F est négative, on obtient que

$$\pm\lambda_2(\Omega) = \iint_D \det\varphi'(u,v)\,\mathrm{d}u\,\mathrm{d}v\,.$$

On en déduit que si $\det \varphi' > 0$, l'orientation induite de F est positive et que si $\det \varphi' < 0$ l'orientation induite de F est négative.



$\det \varphi' > 0$ $\qquad$ $\det \varphi' < 0$

Fig. 6.8

6.2.8 Le flux à travers une courbe plane

Une écriture légèrement différente du théorème de Green donne la formule suivante :

$$\int_{\mathrm{Fr}\,\Omega} (-Q\,\mathrm{d}x + P\,\mathrm{d}y) = \iint_{\Omega} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y \tag{6.16}$$

où Ω, $\mathrm{Fr}\,\Omega$, P et Q sont comme dans l'énoncé de la proposition 6.2.5 ; on a simplement remplacé P par $-Q$ et Q par P dans la formule donnée dans la proposition 6.2.5. On interprète (6.16) comme le ***théorème de la divergence*** dans $\mathbb{R}^2$ pour le champ vectoriel $F = (P, Q)$ en faisant la discussion suivante. L'expression

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}$$

s'appelle la ***divergence*** du champ vectoriel $F = (P, Q)$. On a vu dans le paragraphe 6.2.4 que le vecteur normal ν dirigé vers l'extérieur aux points de $\mathrm{Fr}\,\Omega$ est donné par la formule

$$\nu = \frac{1}{\|\gamma'\|} (\gamma_2', -\gamma_1'),$$

où γ est le paramétrage positif de $\mathrm{Fr}\,\Omega$; ainsi, si $F = (P, Q)$,

$$\int_{\gamma} F \cdot \nu\,\mathrm{d}s = \int_{\mathrm{Fr}\,\Omega} (-Q\,\mathrm{d}x + P\,\mathrm{d}y).$$

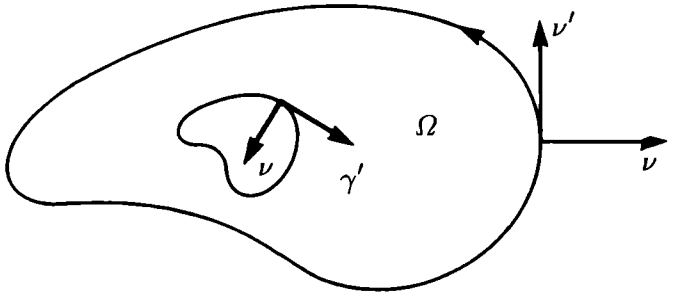


Fig. 6.9

L'expression

$$\int_{\gamma} F \cdot \nu\,\mathrm{d}s$$

s'appelle le ***flux*** de F à travers $\mathrm{Fr}\,\Omega$ (paramétrée par γ) vers l'extérieur de Ω. La formule (6.16) se lit maintenant dans les termes suivants : *le flux du champ $F = (P, Q)$ à travers $\mathrm{Fr}\,\Omega$ vers l'extérieur de Ω est égal à la divergence totale de F dans Ω.*

Cette reformulation ne contient rien de plus mathématique que le théorème de Green donné dans la proposition 6.2.5 ; mais elle a un contenu intui-

tif provenant des différentes interprétations du champ F. Si l'on considère F comme le champ de vitesse d'un écoulement bidimensionnel, on interprète le fait que $\operatorname{div} F = 0$ comme l'incompressibilité du fluide ; (6.16) dit alors simplement que le flux d'un tel champ à travers une courbe fermée est nul ; ceci signifie physiquement que la quantité du fluide qui entre dans Ω par une partie de $\operatorname{Fr}\Omega$ est compensée par sa sortie par une autre partie de $\operatorname{Fr}\Omega$. Si $\operatorname{div} F \neq 0$ l'interprétation est double : soit le fluide n'est pas incompressible, soit il y a une source de perte ou de gain du fluide située à l'intérieur d'un des domaines D_i, $i = 2, \ldots, n$, où $\Omega = D_1 \setminus (\overline{D}_2 \cup \cdots \cup \overline{D}_n)$.

6.2.9 Remarques sur la validité du théorème de Green

Dans le paragraphe 6.1.16, on a donné l'exemple suivant d'un champ vectoriel $F = (P, Q)$ de classe C^∞ dans $\Omega = \mathbb{R}^2 \setminus \{(0,0)\}$:

$$P(x,y) = \frac{-y}{x^2+y^2}, \quad Q(x,y) = \frac{x}{x^2+y^2} ;$$

on a vu que

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0 \quad \text{dans } \Omega$$

mais que

$$\int_{\gamma_r} (P\,\mathrm{d}x + Q\,\mathrm{d}y) = 2\pi$$

pour tout chemin simple circulaire γ_r centré à $(0,0)$ de rayon r et orienté positivement. Le chemin γ_r paramétrise bien le cercle C_r d'équation $x^2 + y^2 = r^2$ qui est la frontière du domaine D_r, le disque de centre $(0,0)$ et de rayon $r > 0$; mais du fait que P et Q ne sont pas dans $C^1(\overline{D}_r)$, on ne peut pas appliquer la formule (6.14) qui est, en fait, en défaut ici. Mais (6.14) est néanmoins utile, même dans cette situation ; soit γ un lacet simple, C^1 par morceaux, ne passant pas par $(0,0)$ et Ω son domaine intérieur ; démontrons que, avec P, Q comme ci-dessus,

$$\int_{\gamma} (P\,\mathrm{d}x + Q\,\mathrm{d}y) = 0 \text{ ou } 2\pi$$

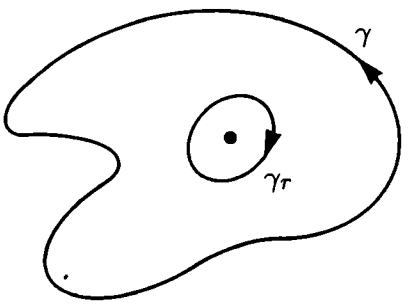


Fig. 6.10

selon que $(0,0) \notin \Omega$ ou $(0,0) \in \Omega$. En effet, si $(0,0) \notin \Omega$, on peut déduire le résultat directement de (6.14) ; si $(0,0) \in \Omega$, on prend un r suffisamment petit pour que $\overline{D}_r \subset \Omega$; alors, selon (6.14) appliqué au domaine $\Omega \setminus \overline{D}_r$, on aura

$$0 = \int_\gamma (P\,\mathrm{d}x + Q\,\mathrm{d}y) - \int_{\gamma_r} (P\,\mathrm{d}x + Q\,\mathrm{d}y)\,,$$

ce qui donne le résultat annoncé. Le procédé utilisé dans cet exemple sera utile pour d'autres calculs. Il ne faut pas conclure que

$$\int_\gamma (P\,\mathrm{d}x + Q\,\mathrm{d}y)$$

sera forcément $\neq 0$ si (P,Q) possède une singularité dans le domaine intérieur défini par la courbe tracée par γ ; par exemple

$$P(x,y) = \frac{x}{r^3}\,, \quad Q(x,y) = \frac{y}{r^3}\,, \quad r = \sqrt{x^2+y^2}$$

donne un champ $F = (P,Q)$ qui possède une singularité à $(0,0)$; mais du fait que $F = \nabla u$ où $u = -1/r$, on conclut que

$$\int_\gamma (P\,\mathrm{d}x + Q\,\mathrm{d}y) = \int_\gamma F \cdot \mathrm{d}\vec{s} = 0$$

dès que γ est un lacet, C^1 par morceaux, qui ne passe pas par $(0,0)$.

Dans le paragraphe suivant, nous donnons une condition géométrique simple sur un domaine Ω qui garantit que $\partial Q/\partial x = \partial P/\partial y$ dans Ω entraîne que le champ $F = (P,Q)$ provient d'un potentiel.

6.2.10 Une condition géométrique

Supposons que Ω est un domaine de $\mathbb{R}^2$ qui possède la propriété suivante (P) : *toute courbe de Jordan polygonale dans Ω est telle que son domaine intérieur est contenu dans Ω.*

Il est facile de voir que dans un domaine $\Omega \subset \mathbb{R}^2$ jouissant de la propriété (P), un champ vectoriel $F = (F_1, F_2)$ de classe C^1 dans Ω est un champ dérivant d'un potentiel dès que $\mathrm{D}_1 F_2 = \mathrm{D}_2 F_1$; en effet, selon le théorème de Green,

$$\int_\gamma F \cdot \mathrm{d}\vec{s} = \iint_D (\mathrm{D}_1 F_2 - \mathrm{D}_2 F_1)\,\mathrm{d}x\,\mathrm{d}y = 0$$

pour tout lacet polygonal γ dans Ω orienté positivement, car le domaine intérieur D à la courbe de Jordan définie par γ est contenu dans Ω ; selon le paragraphe 6.1.12, ceci entraîne que F est conservatif.

La nature géométrique de la propriété (P) est facile à comprendre ; elle dit essentiellement qu'un domaine Ω dans $\mathbb{R}^2$ ayant la propriété (P) ne possède pas de trous. Par exemple, si Ω est un ensemble du type $\mathbb{R}^2 \setminus \{p\}$ ou plus généralement du type $\mathbb{R}^2 \setminus \{p_1, \ldots, p_n\}$, il est alors clair que Ω ne possède pas la propriété (P) ; de même un ensemble Ω annulaire comme

$$\Omega = \{(x,y) : 1 < x^2 + y^2 < 2\}$$

ne possède pas la propriété (P). Par ailleurs, l'exemple vu dans les paragraphes 6.1.16 et 6.2.9 concerne justement l'ensemble $\mathbb{R}^2 \setminus \{(0,0)\}$ ne jouissant pas de la propriété (P).

On peut démontrer que les domaines de $\mathbb{R}^2$ qui ont la propriété (P) sont C^∞-difféomorphes à $\mathbb{R}^2$ (ou à un disque ouvert de $\mathbb{R}^2$). Cette démonstration sera discutée ultérieurement dans la partie du volume 2 consacrée à l'analyse complexe sous la rubrique « Applications conformes ».

Une propriété équivalente à (P) pour un domaine dans $\mathbb{R}^2$ est celle de ***simple connexité*** ; nous formulons celle-ci pour un sous-ensemble quelconque $A \subset \mathbb{R}^N$. On dit que A est ***simplement connexe*** si :

(i) A est connexe par arcs,

(ii) tout lacet γ dans A est ***homotope*** dans A à un lacet constant ce qui veut dire que si $\gamma : [a, b] \to A$, il existe alors une application continue $\Gamma : [a, b] \times [0, 1] \to A$ telle que

$$\Gamma(t,0) = \gamma(t), \quad \Gamma(t,1) = p, \quad a \leqslant t \leqslant b,$$
$$\Gamma(a,u) = \Gamma(b,u), \quad 0 \leqslant u \leqslant 1,$$

où p est un point constant de A ; l'application Γ s'appelle une ***homotopie*** du lacet γ au lacet constant p.

Si nous posons $\gamma_u(t) = \Gamma(t,u)$, on voit que $\gamma_0 = \gamma$, $\gamma_1 \equiv p$ et que, pour chaque $u \in [0, 1]$, γ_u est un lacet dans A ; ainsi, l'homotopie Γ donne une « déformation continue » du lacet γ telle que γ devient un lacet « trivial », c'est-à-dire constant.

Pour nous, l'intérêt de la notion de simple connexité réside dans la validité de la proposition suivante : *soit Ω un domaine simplement connexe dans $\mathbb{R}^N$; un champ vectoriel $F = (F_1, \ldots, F_N) \in C^1(\Omega; \mathbb{R}^N)$ est conservatif si et seulement si $D_j F_i = D_i F_j$, $1 \leqslant j \leqslant N$, dans Ω.*

Nous démontrons cette proposition dans le chapitre 7 dans le contexte des formes différentielles. Nous verrons aussi que les domaines étoilés dans $\mathbb{R}^N$ sont simplement connexes.

Nous avons déjà signalé que pour un domaine $\Omega \subset \mathbb{R}^2$, la propriété (P) est équivalente à la simple connexité. En fait, dans $\mathbb{R}^2$, les domaines simplement connexes peuvent être caractérisés par plusieurs autres conditions géométriques ; voici une condition facile à vérifier : *un domaine Ω dans $\mathbb{R}^2$ est simplement connexe si et seulement si Ω^c ne possède pas de composante connexe* (cf. § 1.8.3 (4)) *bornée ; si le domaine Ω est borné, sa simple connexité est*

équivalente à la connexité de $\Omega^c = \mathbb{R}^2 \setminus \Omega$. Il y a aussi une série de conditions analytiques équivalentes à la simple connexité d'un domaine dans $\mathbb{R}^2$ que nous rencontrons dans l'analyse complexe. Pour une discussion détaillée contenant les démonstrations voir Burckel [7], en particulier page 344.

6.2.11 Exercices

1 Calculer l'aire des ensembles limités par les courbes fermées indiquées (a, b, etc. sont > 0) :

(i) ellipse : $x = a\cos t$, $y = b\sin t$, $0 \leqslant t \leqslant 2\pi$.

(ii) une arche de cycloïde : $x = a(t - \sin t)$, $y = a(1 - \cos t)$, $0 \leqslant t \leqslant 2\pi$ et l'axe des x ;

(iii) boucle du folium de Descartes : $x^3 + y^3 - 3axy = 0$; elle est paramétrée par $x = 3at/(1+t^3)$, $y = 3at^2/(1+t^3)$, $t \geqslant 0$;

(iv) une boucle de la lemmiscate : $r^2 = a^2 \cos 2\theta$ (en coordonnées polaires) ;

(v) cardioïde : $r = a(1 + \cos\theta)$ (en coordonnées polaires) ;

(vi) astroïde : $x = a\cos^3 t$, $y = a\sin^3 t$, $0 \leqslant t \leqslant 2\pi$.

2 Démontrer le théorème de Green pour un cercle centré à l'origine en suivant la méthode du paragraphe 6.2.2.

3 Calculer les intégrales suivantes en utilisant le théorème de Green :

(i) $\displaystyle\int_C \{(2x - y^3)\,\mathrm{d}x + (x^3 + 3y^2)\,\mathrm{d}y\}$,
C le cercle de centre $(0,0)$ et de rayon unité, orienté positivement ;

(ii) $\displaystyle\int_\gamma (\sin y\,\mathrm{d}x + \cos y\,\mathrm{d}y)$,
γ le lacet simple polygonal $(0,0) \to (\pi,0) \to (\pi,\pi/2) \to (0,\pi/2) \to (0,0)$;

(iii) $\displaystyle\int_\gamma \mathrm{e}^{-x}(\sin y\,\mathrm{d}x + \cos y\,\mathrm{d}y)$, γ comme dans (ii) ;

(iv) $\displaystyle\int_\gamma (x^2\,\mathrm{d}x + xy\,\mathrm{d}y)$,
γ le lacet simple polygonal $(0,0) \to (1,0) \to (1,1) \to (0,1) \to (0,0)$;

(v) $\displaystyle\int_C (2xy\,\mathrm{d}x - 3xy\,\mathrm{d}y)$,
C est le carré de sommets $(3,1) \to (5,1) \to (5,3) \to (3,3)$ orienté négativement.

4 Vérifier, en suivant la discussion du paragraphe 6.2.4, que

$$\gamma(t) = (\cos t, \sin t), \quad 0 \leqslant t \leqslant 2\pi$$

est bien un paramétrage positif du cercle unité d'équation $x^2 + y^2 = 1$ dans $\mathbb{R}^2$.

5 Calculer $\int_\gamma (f\,dx + g\,dy)$ où

$$f(x,y) = \frac{e^x}{x^2+y^2}\,(x\sin y - y\cos y)\,, \quad g(x,y) = \frac{e^x}{x^2+y^2}\,(x\cos y + y\sin y)$$

et γ est un lacet simple, C^1 par morceaux, dans $\mathbb{R}^2 \setminus \{(0,0)\}$, orienté positivement.

REMARQUE. Vérifier d'abord que $g'_x = f'_y$; cet exercice devient plus compréhensible en considérant la fonction complexe donnée par e^z/z.

6 Calculer $\int_\gamma (P\,dx + Q\,dy)$ où

$$P(x,y) = \frac{-y}{(x-1)^2+y^2}\,, \quad Q(x,y) = \frac{x(x-1)+y^2}{(x-1)^2+y^2}$$

et γ est un lacet simple, C^1 par morceaux, dans $\mathbb{R}^2 \setminus \{(0,0)\}$, orienté positivement.

7 Soit

$$I_i = \int_{\gamma_i} \{(x^2+y^2)\,dx + (x^2-y^2)\,dy\}\,, \quad i = 1,\,2\,,$$

où γ_1 est le segment droit de $(0,0)$ à $(1,1)$ et γ_2 le morceau de parabole d'équation $y = x^2$ de $(0,0)$ à $(1,1)$; démontrer que $I_2 > I_1$.

8 Démontrer que

$$\int_C \{(\cos xy + yx^3 + e^y)\,dx + (x\,e^y + xy^3)\,dy\} = 0$$

où C est une courbe de Jordan, C^1 par morceaux, orientée positivement par rapport à son domaine intérieur D, et telle que $(-x,-y) \in D$ dès que $(x,y) \in D$.

9 Calculer

$$\int_C \{y^2\,dx + (2xy + x)\,dy\}$$

où C est la frontière du domaine

$$D = \{(x,y) : y - x^2 > 0\,,\ 0 < y < 1\}\,,$$

C étant orienté positivement par rapport à D.

10 Calculer directement

$$\int_\gamma \frac{\mathrm{d}y}{x^2+y^2}$$

où γ est le lacet simple polygonal $(1,1) \to (-1,1) \to (-1,-1) \to (1,-1) \to (1,1)$; essayer d'obtenir le résultat par une application appropriée du théorème de Green.

11 Calculer les intégrales suivantes où γ est un chemin quelconque, C^1 par morceaux, reliant les points indiqués dans les ensembles ouverts précisés :

(i) $\displaystyle\int_\gamma \frac{-y\,\mathrm{d}x + x\,\mathrm{d}y}{x^2+y^2}$, de $(1,1)$ à $(0,1)$ dans $\mathbb{R}^2 \setminus \mathbb{R}_-$;

(ii) $\displaystyle\int_\gamma \frac{x\,\mathrm{d}x + y\,\mathrm{d}y}{x^2+y^2}$, de $(1,0)$ à $(1,1)$ dans $\mathbb{R}^2 \setminus \{(0,0)\}$;

(iii) $\displaystyle\int_\gamma \frac{x\,\mathrm{d}x + y\,\mathrm{d}y}{\sqrt{x^2+y^2}}$, de $(1,0)$ à $(0,2)$ dans $\mathbb{R}^2 \setminus \{(0,0)\}$;

(iv) $\displaystyle\int_\gamma \frac{-yz\,\mathrm{d}x + \{xz + (x-yz)^2\}\,\mathrm{d}y + \{xy + (x-yz)^2\}\,\mathrm{d}z}{(x-yz)^2}$,

de $(1,0,0)$ à $(2,1,1)$ dans $\{(x,y,z) : x - yz \neq 0\}$;

(v) $\displaystyle\int_\gamma f\Big(\frac{yz}{x}\Big) \cdot \frac{-yz\,\mathrm{d}x + zx\,\mathrm{d}y + yx\,\mathrm{d}z}{x^2}$,

de $(1,0,0)$ à $(1,1,1)$ dans $\{(x,y,z) : x \neq 0\}$.

12 Soit $F = (F_1, F_2)$ un champ vectoriel de classe C^1 dans $\Omega = \mathbb{R}^2 \setminus \{p_1, p_2\}$; on suppose que $\mathrm{D}_1 F_2 = \mathrm{D}_2 F_1$ dans Ω. Soit $I_i = \int_{C_i} F \cdot \mathrm{d}\vec{s}$, $i = 1, 2$, où C_i est un cercle centré en p_i de rayon $r_i < \|p_1 - p_2\|$, orienté positivement. Si γ est un lacet simple, C^1 par morceaux, dans Ω, démontrer que $\int_\gamma F \cdot \mathrm{d}\vec{s} = \alpha_1 I_1 + \alpha_2 I_2$ où α_1, α_2 sont des entiers valant 0, 1 ou -1.

13 Soit $F = (F_1, F_2)$ un champ vectoriel de classe C_1 dans $\Omega = \mathbb{R}^2 \setminus \{(0,0)\}$. Démontrer que F est un champ conservatif si et seulement si :

(i) $\mathrm{D}_1 F_2 = \mathrm{D}_2 F_1$ dans Ω,

(ii) $\int_C F \cdot \mathrm{d}\vec{s} = 0$ où C est le cercle d'équation $x_1^2 + x_2^2 = 1$, orienté positivement.

6.3 INTÉGRALES DE SURFACE DANS $\mathbb{R}^3$

Dans cette section, nous définissons les intégrales des fonctions scalaires définies sur une surface régulière dans $\mathbb{R}^3$ et le flux d'un champ vectoriel à travers une telle surface orientable. Une surface régulière dans $\mathbb{R}^3$ est, par définition, une sous-variété différentielle de dimension 2 dans $\mathbb{R}^3$, de classe C^p, $p \geqslant 1$ (cf. § 4.4.2) ; mais, dans cette section ainsi que dans le reste de ce chapitre, les calculs seront limités aux morceaux de surfaces régulières dans $\mathbb{R}^3$. La théorie globale sera reprise dans le chapitre 7 pour les sous-variétés différentielles de classe C^p, de dimension quelconque, dans $\mathbb{R}^N$; elle est seulement esquissée dans ce chapitre.

6.3.1 Morceaux de surfaces régulières dans $\mathbb{R}^3$

Une partie $S \subset \mathbb{R}^3$ s'appelle un ***morceau de surface régulière*** de classe C^1 (brièvement : ***morceau de surface***) s'il existe un domaine U dans $\mathbb{R}^2$ et une application $\psi : U \to \mathbb{R}^3$ telle que :

(i) $\psi(U) = S$ et ψ est un homéomorphisme entre U et S ;
(ii) ψ est de classe C^1 ;
(iii) pour tout $u \in U$, l'application linéaire $\psi'(u) : \mathbb{R}^2 \to \mathbb{R}^3$ est de rang 2.

On dit que S est ***paramétré*** par (ψ, U) ou, plus brièvement, par ψ ; la variable u s'appelle le ***paramètre*** ; S est de ***classe*** C^p ($p \geqslant 1$) si ψ est de classe C^p ; on dira aussi que S est donné par une ***représentation paramétrique*** (ψ, U).

Si $u = (u_1, u_2)$, $\psi = (\psi_1, \psi_2, \psi_3)$, la condition (iii) peut être reformulée comme suit :

(iii′) la matrice $[\mathrm{D}_j\psi_i(u)]$ $(i = 1, 2, 3, j = 1, 2)$ est de rang 2 pour tout $u \in U$;
(iii″) pour tout $u \in U$, $\mathrm{D}_1\psi(u) \times \mathrm{D}_2\psi(u) \neq 0$ où

$$\mathrm{D}_j\psi(u) = \big(\mathrm{D}_j\psi_1(u), \mathrm{D}_j\psi_2(u), \mathrm{D}_j\psi_3(u)\big)\,, \quad j = 1, 2\,,$$

et $\times$ signifie le produit vectoriel dans $\mathbb{R}^3$. L'équivalence entre (iii), (iii′) et (iii″) est immédiate.

Un calcul montre que

$$\mathrm{D}_1\psi(u) \times \mathrm{D}_2\psi(u) = \big(a_1(u), a_2(u), a_3(u)\big)$$

où

$$a_1 = \frac{\partial(\psi_2, \psi_3)}{\partial(u_1, u_2)}\,, \quad a_2 = \frac{\partial(\psi_3, \psi_1)}{\partial(u_1, u_2)}\,, \quad a_3 = \frac{\partial(\psi_1, \psi_2)}{\partial(u_1, u_2)}\,.$$

Selon (iii″), pour chaque $u \in U$ une des quantités $a_j(u)$ est non nulle. Fixons un $v = (v_1, v_2)$ dans U et supposons que $a_3(v) \neq 0$; soit $\psi(v) = p$, $p = (p_1, p_2, p_3)$.

Selon le théorème d'inversion locale (cf. chap. 4), les équations

$$x_1 = \psi_1(u_1, u_2), \quad x_2 = \psi_2(u_1, u_2)$$

peuvent être résolues uniquement dans un voisinage ouvert V de v, $V \subset U$, dans la forme

$$u_1 = \varphi_1(x_1, x_2), \quad u_2 = \varphi_2(x_1, x_2)$$

où $\varphi = (\varphi_1, \varphi_2)$ est une fonction de classe C^1 dans un voisinage ouvert W de (p_1, p_2) avec $\varphi(p_1, p_2) = (v_1, v_2)$; de plus, $\varphi : W \to V$ est un C^1-difféomorphisme entre W et V. Posons

$$f(x_1, x_2) = (\psi \circ \varphi)(x_1, x_2), \quad (x_1, x_2) \in W$$

et

$$T = \{(x_1, x_2, x_3) : (x_1, x_2) \in W, \ x_3 = f(x_1, x_2)\}.$$

Alors $T = \psi(V) \subset S$; T est une partie de S formant un morceau de surface paramétré par (ξ, W) où

$$\xi(x_1, x_2) = \big(x_1, x_2, f(x_1, x_2)\big), \quad (x_1, x_2) \in W.$$

Notons que ψ étant un homéomorphisme entre U et S, $T = \psi(V)$ est S-ouvert de sorte que $T = S \cap A$ où A est un ensemble ouvert de $\mathbb{R}^3$. On dit que le morceau de surface T est représenté par une ***équation cartésienne***

$$x_3 = f(x_1, x_2), \quad (x_1, x_2) \in W, \tag{6.17}$$

où W est une partie ouverte de $\mathbb{R}^2$ et $f \in C^1(W)$. La présentation (6.17) est obtenue en prenant pour hypothèse que $a_3(v) \neq 0$, avec $\psi(v) = p = (p_1, p_2, p_3)$, W étant un voisinage ouvert de (p_1, p_2) et $f(p_1, p_2) = p_3$; nous dirons que S est représenté par (6.17) dans un voisinage ouvert du point p. Il est clair que si $a_1(v) \neq 0$, S sera représenté dans un voisinage ouvert de p par une équation cartésienne du type

$$x_1 = f(x_2, x_3), \quad (x_2, x_3) \in W$$

et si $a_2(v) \neq 0$, on aura une équation cartésienne du type

$$x_2 = f(x_1, x_3), \quad (x_1, x_3) \in W.$$

La discussion précédente peut être résumée comme suit : *tout morceau de surface régulière est localement représentable par une équation cartésienne.*

Notons qu'une représentation en équation cartésienne est un cas particulier d'une représentation paramétrique ; aussi (6.17) correspond à la représentation paramétrique (ξ, W) de T où

$$\xi(x_1, x_2) = \big(x_1, x_2, f(x_1, x_2)\big), \quad (x_1, x_2) \in W.$$

Une telle représentation a pour avantage que les conditions (i), (ii) et (iii) sont immédiatement vérifiées dès que $f \in C^1(W)$. En effet, ξ est alors clairement de classe C^1 et $\xi^{-1} : T \to W$ est évidemment continue, car *elle est restriction* à T de la projection $(x_1, x_2, x_3) \mapsto (x_1, x_2)$; puisque

$$\mathrm{D}_1\xi = \big(1, 0, \mathrm{D}_1 f\big)\,, \quad \mathrm{D}_2\xi = \big(0, 1, \mathrm{D}_2 f\big)\,,$$

on voit immédiatement que ξ' est de rang 2 partout. Pour la suite, nous calculons

$$\mathrm{D}_1\xi \times \mathrm{D}_2\xi = (-\mathrm{D}_1 f, -\mathrm{D}_2 f, 1)$$

ce qui est une nouvelle confirmation que ξ' est de rang 2 partout.

Notons un résultat technique important qui découle de la discussion amenant vers (6.17). *Si (ψ, U) est une représentation paramétrique du morceau de surface régulière S dans $\mathbb{R}^3$ alors $\psi^{-1} : S \to U$ est de classe C^1*. Notons que S n'est certainement pas une partie ouverte de $\mathbb{R}^3$ et le sens de l'appartenance à la classe C^1 de ψ^{-1} est celui donné à cette phrase dans le paragraphe 3.9.1 ; en effet, pour chaque $p \in S$, il existe un voisinage ouvert A de p (A étant un ensemble ouvert de $\mathbb{R}^3$) tel que $\psi^{-1}|_{S \cap A}$ est égale à $\varphi \circ \pi$ où $\pi(x_1, x_2, x_3) = (x_1, x_2)$ et où φ, A sont définis précédemment (pourvu que $a_3(v) \neq 0$, p étant $\psi(v)$) ; donc $\psi^{-1}|_{S \cap A}$ possède un prolongement de classe C^1 à l'ensemble ouvert A car $\varphi \circ \pi$ est clairement de classe C^1 dans A ; une discussion analogue étant valable si $a_1(v) \neq 0$ ou si $a_2(v) \neq 0$, notre affirmation concernant le caractère C^1 de ψ^{-1} est établie. Puisque le calcul de $\psi^{-1\prime}(p)$ nécessite seulement un prolongement de ψ^{-1} à un voisinage ouvert de p, nous ne discuterons pas ici la possibilité du prolongement de ψ^{-1} à un ensemble ouvert contenant S ; une démonstration de ce fait est contenue dans le chapitre 7.

Notons que $\psi^{-1}\big(\psi(u)\big) = u$ entraîne que

$$\psi^{-1\prime}\big(\psi(u)\big)\psi'(u) = I$$

où I est l'application identité de $\mathbb{R}^2$ dans $\mathbb{R}^2$; si $\psi(v) = a$, on en déduit que $\psi^{-1\prime}(a)(T_a) = \mathbb{R}^2$ où $T_a = \psi'(v)(\mathbb{R}^2)$ est l'***espace tangent*** à S au point a (§ 6.3.4) ; puisque la dimension de T_a est égale à la dimension de $\mathbb{R}^2$ on arrive à la *conclusion importante suivante : si $\xi \in T_a$, $\xi \neq 0$, alors $\psi^{-1\prime}(a)\xi \neq 0$.*

Une explication intuitive pour motiver la définition proposée pour un morceau de surface est la suivante : imaginons une matière plastique bidimensionnelle représentée par $U \subset \mathbb{R}^2$; le morceau de surface S est obtenu par une déformation régulière ψ de U.

6.3.2 Remarques concernant la propriété homéomorphique de ψ

Les conditions (ii) et (iii) dans la définition d'un morceau de surface sont assez simples à vérifier ; la vérification de la condition (i) exige (en vue de la condition (ii)) que l'on établisse les points suivants :

(a) ψ est injective ;
(b) ψ^{-1} est continue.

On a vu que ces deux points sont automatiquement vérifiés dans la représentation du morceau de surface par une équation cartésienne ; par ailleurs, cette représentation à l'avantage de satisfaire aussi aux conditions (ii) et (iii) dès que la fonction f dans sa représentation est de classe C^1. Pour une représentation paramétrique générale (ψ, U), il se peut que (a) soit vérifiée sans que (b) le soit ; néanmoins, *si $\psi : U \to \mathbb{R}^3$ est injective et remplit les conditions (ii) et (iii) du paragraphe 6.3.1 alors la restriction de ψ à une partie ouverte bornée V telle que $\overline{V} \subset U$ définit un morceau de surface $\psi(V)$*. En effet, $\overline{V}$ étant compact, l'injectivité de ψ entraîne que $\psi|_{\overline{V}}$ est un homéomorphisme entre $\overline{V}$ et $\psi(\overline{V})$ (prop. 1.7.3(iii)).

A cause de ce qui précède, il nous sera rarement nécessaire de vérifier en détail que ψ est un homéomorphisme. Néanmoins, nous donnons une proposition qui est utile dans ce contexte : *soit $\psi : \overline{U} \to \mathbb{R}^M$ une application continue telle que U est un ensemble ouvert borné dans $\mathbb{R}^N$ avec $\psi|_U$ injective et*

$$\psi(\operatorname{Fr} U) \cap \psi(U) = \emptyset \, ;$$

alors $\psi|_U$ est un homéomorphisme entre U et $\psi(U)$.

Pour la démonstration, il suffit d'établir que si $p_n \in U$, $q_n = \psi(p_n)$, $n = 1, 2, \ldots$, et si $q_n \to q$ lorsque $n \to \infty$ avec $q \in \psi(U)$ alors $p_n \to p$ lorsque $n \to \infty$ avec $p \in U$ et $\psi(p) = q$; en effet, comme $\overline{U}$ est compact, il existe une sous-suite $\{p_{n_k}\}$ telle que $p_{n_k} \to \widetilde{p} \in \overline{U}$ $(k \to \infty)$; mais alors $\psi(\widetilde{p}) = q$, ce qui exclut que $\widetilde{p} \in \operatorname{Fr} U$ car $\psi(\operatorname{Fr} U)$ et $\psi(U)$ sont disjoints ; ainsi, $\widetilde{p} \in U$; à cause de l'injectivité de $\psi|_U$, on voit que toute sous-suite convergente $\{p_{n_k}\}$ converge vers un même élément $\widetilde{p} = p$ de U de sorte que $p_n \to p$ lorsque $n \to \infty$.

6.3.3 Intégrales de surface

Soit S un morceau de surface régulière dans $\mathbb{R}^3$ paramétré par (ψ, U). Si $f : S \to \mathbb{R}$, on *définit* $\int_S f \, d\sigma$ par la formule

$$\int_S f \, d\sigma = \iint_U f(\psi(u)) \left\| D_1\psi(u) \times D_2\psi(u) \right\| du_1 \, du_2 \tag{6.18}$$

pourvu que l'intégrale à droite de (6.18) existe ; le côté gauche de (6.18) s'appelle l'***intégrale de*** f sur le morceau de surface S. Si $f = 1_A$, $A \subset S$, dans (6.18) on aura

$$\sigma(A) = \iint_{\psi^{-1}(A)} \left\| D_1\psi \times D_2\psi \right\| du_1 \, du_2 \tag{6.19}$$

où l'on a posé

$$\sigma(A) = \int_S 1_A \, d\sigma \, .$$

On dit que $\sigma(A)$ est la ***mesure superficielle*** ou l'***aire*** de $A \subset S$. Dans ces considérations, nous avons tacitement admis que A est une partie borélienne de S et $f : S \to \mathbb{R}$ est une fonction borélienne ; nous admettons ceci dans

la suite sans nécessairement le rappeler explicitement. En munissant S de sa tribu borélienne Σ_S, la formule (6.19) définit une mesure positive sur (S, Σ_S) ; la formule (6.18) dit que $\int_S f \, d\sigma$ est simplement l'intégrale de $f : S \to \mathbb{R}$ par rapport à cette mesure positive σ. Souvent, on utilise l'écriture symbolique suivante :

$$d\sigma = \left\| D_1\psi \times D_2\psi \right\| du_1 \, du_2 \,. \tag{6.19'}$$

On dit aussi que l'***élément d'aire*** $d\sigma$ sur S est défini par cette dernière formule. Une discussion géométrique de celle-ci est donnée dans la suite.

Si S est défini par une équation cartésienne

$$x_3 = f(x_1, x_2) \,, \quad (x_1, x_2) \in W \,,$$

avec $f \in C^1(W)$, W un domaine dans $\mathbb{R}^2$, un calcul simple partant de

$$\psi(x_1, x_2) = \big(x_1, \, x_2, \, f(x_1, x_2)\big)$$

donne que

$$d\sigma = \left\{ D_1^2 f + D_2^2 f + 1 \right\}^{1/2} dx_1 \, dx_2 \,. \tag{6.20}$$

Revenons au cas de S donné par un paramétrage (ψ, U) et posons

$$A(u) = \left[D_j \psi_i(u) \right]$$

pour la matrice de type (3,2) qui représente $\psi'(u)$ dans les bases canoniques où $\psi = (\psi_1, \psi_2, \psi_3)$. Un calcul simple montre que

$$^tAA = \begin{bmatrix} E & F \\ F & G \end{bmatrix}$$

où

$$E = \left\| D_1\psi \right\|^2 , \quad F = D_1\psi \cdot D_2\psi \,, \quad G = \left\| D_2\psi \right\|^2 ,$$

en supprimant l'argument $u \in U$ partout ; on a alors la formule suivante :

$$\left\| D_1\psi \times D_2\psi \right\| = \left\{ \det(^tAA) \right\}^{1/2} = \sqrt{EG - F^2} \,. \tag{6.21}$$

En effet, de la formule

$$\left\| a \times b \right\|^2 + (a.b)^2 = \left\| a \right\|^2 \left\| b \right\|^2$$

appliquée à $a = D_1\psi$, $b = D_2\psi$, on obtient

$$\left\| D_1\psi \times D_2\psi \right\|^2 = EG - F^2 = \det(^tAA) \,,$$

ce qui donne (6.21) et justifie l'écriture

$$d\sigma = \sqrt{EG - F^2} \, du_1 \, du_2 = \left\{ \det(^tAA) \right\}^{1/2} du_1 \, du_2 \,.$$

L'écriture (6.21) possède l'avantage d'être généralisable aux variétés de dimension d quelconque ; en effet, la formule pour la mesure d-dimensionnelle d'un parallélépipède de côtés donnés par des vecteurs $\xi_1, \ldots, \xi_d$ dans $\mathbb{R}^N$ était précisée dans (5.25) dont le cas particulier $d = 2$ concerne la formule bien connue de l'aire d'un parallélogramme de côtés donnés par les vecteurs a, b dans $\mathbb{R}^2$ (soit $\|a \times b\|$).

L'interprétation géométrique pour la formule

$$\mathrm{d}\sigma = \left\|\mathrm{D}_1\psi \times \mathrm{D}_2\psi\right\| \mathrm{d}u_1\, \mathrm{d}u_2$$

est la suivante : un rectangle infinitésimal de côtés de longueur $\mathrm{d}u_1$, $\mathrm{d}u_2$ au point $u = (u_1, u_2)$ dans U devient, sous l'application ψ, un parallélogramme infinitésimal dans S de côtés donnés par les vecteurs $\mathrm{D}_1\psi\, \mathrm{d}u_1$, $\mathrm{D}_2\psi\, \mathrm{d}u_2$ dont l'aire est

$$\left\|\mathrm{D}_1\psi \times \mathrm{D}_2\psi\right\| \mathrm{d}u_1\, \mathrm{d}u_2\,.$$

Cette discussion géométrique essaie d'être suggestive ; nous n'attribuons aucune valeur mathématique à celle-ci. Quelques autres considérations sur la définition d'aire superficielle sont données dans le paragraphe 6.3.10.

Les quantités E, F et G qui apparaissent dans (6.1) jouent un rôle fondamental dans l'étude de la surface S ; elles définissent « la première forme fondamentale » de la surface S qui reflète toutes les propriétés « métriques » (comme distance entre les points et comme angle entre deux courbes) dans S. L'étude de ces propriétés est un sujet de la géométrie différentielle que nous n'aborderons pas dans ce livre.

6.3.4 Normale unitaire, orientation, flux

Soit S un morceau de surface régulière dans $\mathbb{R}^3$ paramétré par (ψ, U). L'***espace tangent*** à S au point $a = \psi(v)$ est, par définition, le sous-espace vectoriel T_a de $\mathbb{R}^3$ défini par $\psi'(v)(\mathbb{R}^2)$; puisque $\psi'(v)$ est de rang 2, T_a est de dimension 2 ; une base pour T_a est donnée par $\{\mathrm{D}_1\psi(v), \mathrm{D}_2\psi(v)\}$. Il est clair que le sous-espace vectoriel de $\mathbb{R}^3$, orthogonal à T_a (soit $T_a^\perp$) est de dimension 1 et qu'il est engendré par le vecteur $\mathrm{D}_1\psi(v) \times \mathrm{D}_2\psi(v)$; Les vecteurs de $T_a^\perp$ s'appellent les ***vecteurs normaux*** à S au point $a = \psi(v)$; ces vecteurs normaux sont les multiples réels de $\mathrm{D}_1\psi(v) \times \mathrm{D}_2\psi(v)$; les deux ***normales unitaires*** à S au point a, sont, par définition,

$$\pm \frac{\mathrm{D}_1\psi(v) \times \mathrm{D}_2\psi(v)}{\left\|\mathrm{D}_1\psi(v) \times \mathrm{D}_2\psi(v)\right\|}\,.$$

Le choix du signe + pour tout $v \in U$ définit une ***orientation*** pour S et le choix du signe − pour tout $v \in U$ donne l'autre orientation pour S ; ces deux orientations sont dites ***opposées*** l'une à l'autre. Nous fixerons une orientation par la définition continue de la normale unitaire $\nu(a)$ à chaque point a de S ; ceci revient à choisir partout le même signe dans la formule précédente pour

une normale unitaire ; notons qu'en choisissant le signe + partout, c'est-à-dire en prenant

$$\nu(a) = c\mathrm{D}_1\psi(v) \times \mathrm{D}_2\psi(v)\,, \quad c > 0\,,\ a \in S\,,$$

on a trois vecteurs $\{\mathrm{D}_1\psi(v),\ \mathrm{D}_2\psi(v),\ \nu(a)\}$ qui forment une base positive pour $\mathbb{R}^3$ pour tout point $a \in S$, $a = \psi(v)$ ce qui veut dire, rappelons-le, que

$$\det[\mathrm{D}_1\psi(v),\ \mathrm{D}_2\psi(v),\ \nu(a)] > 0\,;$$

c'est une conséquence du fait que si x, y sont deux vecteurs linéairement indépendants de $\mathbb{R}^3$,

$$\begin{aligned}\det[x,\ y,\ x \times y] &= x \cdot \{y \times (x \times y)\} \\ &= x \cdot \left\{\|y\|^2 x - (x \cdot y)y\right\} = \|x\|^2 \|y\|^2 - (x \cdot y)^2 > 0\,.\end{aligned}$$

Rappelons que le plan tangent (de la géométrie usuelle) à S au point $a = \psi(v)$ est défini par l'équation suivante :

$$(x - a) \cdot \xi = 0$$

où $\xi \neq 0$ est un vecteur de $T_a^{\perp}$; on peut prendre, par exemple, $\xi = \mathrm{D}_1\psi(v) \times \mathrm{D}_2\psi(v)$. Notons la propriété caractéristique des vecteurs de T_a (appelés aussi ***vecteurs tangents*** à S au point $a = \psi(v)$) : *soit $\alpha : [-\varepsilon\,,\ \varepsilon] \to U$, $\varepsilon > 0$, un chemin régulier dans U avec $\alpha(0) = v$; si $\gamma = \psi \circ \alpha$, γ est un chemin régulier dans S tel que $\gamma(0) = a$, $\gamma'(0) \in T_a$. De plus, tout chemin régulier γ dans S est de la forme $\gamma = \psi \circ \alpha$ où α est un chemin régulier dans U.*

En effet, la relation

$$\gamma'(t) = \psi'\big(\alpha(t)\big)\alpha'(t)$$

montre que $\gamma'(t) \neq 0$ car $\alpha'(t) \neq 0$ (α étant un chemin régulier) et $\psi'\big(\alpha(t)\big)$ est de rang 2 ; avec $t = 0$, on obtient que $\gamma'(0) = \psi'(v)\alpha'(0) \in T_a$. D'autre part, si γ est un chemin dans S, en prenant $\alpha = \psi^{-1} \circ \gamma$ on aura un chemin dans U tel que $\gamma = \psi \circ \alpha$; si γ est de classe C^1, $\alpha = \psi^{-1} \circ \gamma$ est de C^1 car ψ^{-1} est de classe C^1 (cf. § 6.3.1) ; de plus, on aura

$$\alpha'(t) = \psi^{-1\prime}\big(\gamma(t)\big)\gamma'(t)\,,$$

ce qui donne que $\alpha'(t) \neq 0$ si $\gamma'(t) \neq 0$ (cf. § 6.3.1), démontrant que α est un chemin régulier si γ est régulier.

Soit $F : S \to \mathbb{R}^3$ un champ vectoriel dans S. Le ***flux de F à travers S orienté*** par un choix de la normale unitaire ν est défini par la formule

$$\int_S F \cdot \nu \,\mathrm{d}\sigma$$

en supposant que cette intégrale existe. De (6.18), on voit que

$$\int_S F \cdot \nu \, d\sigma = \pm \iint_U F(\psi(u)) \cdot \{D_1\psi(u) \times D_2\psi(u)\} \, du_1 \, du_2$$
$$= \pm \iint_U [F \circ \psi, D_1\psi, D_2\psi] \, du$$

où $[a, b, c] = a \cdot (b \times c)$ est le produit mixte des vecteurs a, b et c dans $\mathbb{R}^3$ car

$$F \cdot \nu = \pm F \cdot \left\{ \frac{D_1\psi \times D_2\psi}{\|D_1\psi \times D_2\psi\|} \right\}, \quad d\sigma = \|D_1\psi \times D_2\psi\| \, du_1 \, du_2 .$$

Evidemment, le flux dépend du choix de l'orientation ; en renversant l'orientation, on change le signe du flux.

Si $F(x)$ est interprété comme le vecteur vitesse d'un fluide de densité $\rho(x)$ au point x alors

$$\rho(x)F(x) \cdot \nu(x) \, d\sigma$$

peut être considéré comme la quantité de fluide traversant une partie infinitésimale de la surface S par unité de temps ; l'intégration sur S donne le flux de ρF à travers S qui peut être interprété comme l'écoulement total du fluide à travers S par unité de temps. Des interprétations analogues sont possibles lorsque F représente un champ électrique ou magnétique.

6.3.5 Changement de paramétrage

On a vu qu'un morceau de surface peut être représenté par un paramétrage ou par une équation cartésienne, cette dernière étant, en fait, un paramétrage d'un type particulier. En vérité, un morceau de surface peut être paramétré d'une infinité de manière ; si (ψ, U) est un paramétrage donné pour un morceau de surface S dans $\mathbb{R}^3$ et si $\theta : V \to U$ est un C^1-difféormorphisme de l'ensemble ouvert $V \subset \mathbb{R}^2$ sur U alors $\xi = \psi \circ \theta$ donne un nouveau paramétrage (ξ, V) pour le même morceau de surface S. Inversement, *si* (ψ, U), (ξ, V) *sont deux paramétrages de* S *alors* $\theta = \psi^{-1} \circ \xi$ *est un* C^1*-difféomorphisme de* V *sur* U. En effet, selon la discussion du paragraphe 6.3.1, ψ^{-1} et ξ sont de classe C^1 de sorte que θ est de classe C^1 et $\theta : V \to U$ est un homéomorphisme ; puisque $\theta^{-1} = \xi^{-1} \circ \psi$, le même raisonnement donne que θ^{-1} est aussi de classe C^1 d'où l'affirmation que $\theta : V \to U$ est un C^1-difféomorphisme de V sur U.

Nous vérifions maintenant que *si* $f : S \to \mathbb{R}$ *est une fonction réelle pour laquelle* $\int_S f \, d\sigma$ *existe selon un paramétrage* (ψ, U), *cette intégrale existe également selon le paramétrage* (ξ, V) *et les deux calculs donnent le même résultat*. La démonstration est une conséquence de la formule de changement de variables dans les intégrales doubles et le calcul suivant : soit $\theta : V \to U$ le C^1-difféomorphisme tel que $\xi = \psi \circ \theta$; alors

$$\xi'(v) = \psi'(\theta(v))\theta'(v), \quad v \in V .$$

En posant A, B, C pour les matrices canoniques représentant $\psi'(\theta(v))$, $\xi'(v)$, $\theta'(v)$ respectivement, on aura

$$B = AC\,, \quad {}^tBB = {}^tC({}^tAA)C$$

d'où

$$\det({}^tBB) = \det({}^tAA)\det{}^2(C)\,,$$

ce qui donne, grâce à (6.21), que

$$\|\mathrm{D}_1\xi(v) \times \mathrm{D}_2\xi(v)\| = \|\mathrm{D}_1\psi(\theta(v)) \times \mathrm{D}_2\psi(\theta(v))\| \cdot |\det\theta'(v)|\,.$$

On en déduit, en faisant le changement de variables $u = \theta(v)$, les formules suivantes :

$$\begin{aligned}\int_S f\,\mathrm{d}\sigma &= \int_U f(\psi(u))\|\mathrm{D}_1\psi(u) \times \mathrm{D}_2\psi(u)\|\,\mathrm{d}u \\ &= \int_V f(\psi(\theta(v)))\,\|\mathrm{D}_1\psi(\theta(v)) \times \mathrm{D}_2\psi(\theta(v))\| \cdot |\det\theta'(v)|\,\mathrm{d}v \\ &= \int_V f(\xi(v))\|\mathrm{D}_1\xi(v) \times \mathrm{D}_2\xi(v)\|\,\mathrm{d}v\,.\end{aligned}$$

Ces égalités établissent notre affirmation concernant l'invariance de $\int_S f\,\mathrm{d}\sigma$ sous changement de paramétrage. Dorénavant, on dira que f est ***intégrable sur*** S si $\int_S |f|\,\mathrm{d}\sigma < \infty$ ce qui équivaut, dans notre théorie d'intégration (c'est-à-dire celle de Lebesgue), à l'existence de $\int_S f\,\mathrm{d}\sigma$.

Etablissons aussi que *l'espace tangent T_a pour $a \in S$, $a = \psi(u)$, et indépendant du paramétrage (ψ, U) utilisé*. En effet, si (ξ, V) est un autre paramétrage avec $\xi = \psi \circ \theta$ comme ci-dessus et $\theta(v) = u$ alors la relation

$$\xi'(v) = \psi'(u)\theta'(v)$$

donne que

$$\xi'(v)(\mathbb{R}^2) = \psi'(u)\{\theta'(v)(\mathbb{R}^2)\} = \psi'(u)(\mathbb{R}^2) = T_a$$

car $\theta'(v)$ est une bijection de $\mathbb{R}^2$ sur $\mathbb{R}^2$; ceci établit que la définition de T_a ne dépend pas du choix de paramétrage. Une conséquence immédiate est que *les deux normales unitaires en chaque point sont indépendantes du paramétrage choisi*. Ainsi, le calcul du flux d'un champ vectoriel à travers un morceau de surface est aussi indépendant du paramétrage utilisé, pour autant, naturellement, que l'orientation fixée reste la même lors d'un passage à un nouveau paramétrage.

6.3.6 D'autres notations pour le flux

Soit S un morceau de surface dans $\mathbb{R}^3$ paramétré par (ψ, U) ; fixons un point $u \in U$ avec $\psi(u) = x$ et une normale unitaire $\nu = \nu(x)$ dans la direction de $\mathrm{D}_1\psi \times \mathrm{D}_2\psi$ en ce point x ; en posant

$$\cos\alpha_i(x) = \nu(x) \cdot e_i\,, \quad 1 \leqslant i \leqslant 3\,,$$

où $\alpha_i(x) \in [0, \pi]$, les e_i étant les trois vecteurs de la base canonique de $\mathbb{R}^3$, on voit que

$$\nu(x) = \big(\cos\alpha_1(x), \cos\alpha_2(x), \cos\alpha_3(x)\big) ;$$

α_i est interprété géométriquement comme l'angle entre ν et l'axe des x_i. L'expression $\nu\,\mathrm{d}\sigma$ qui apparaît dans la définition du flux $\int_S F \cdot \nu\,\mathrm{d}\sigma$ d'un champ vectoriel $F : S \to \mathbb{R}^2$ peut être écrit symboliquement comme

$$\nu\,\mathrm{d}\sigma = (\cos\alpha_1, \cos\alpha_2, \cos\alpha_3)\,\mathrm{d}\sigma = (\mathrm{D}_1\psi \times \mathrm{D}_2\psi)\,\mathrm{d}u\,,$$

car

$$(\cos\alpha_1, \cos\alpha_2, \cos\alpha_3) = \frac{\mathrm{D}_1\psi \times \mathrm{D}_2\psi}{\big\|\mathrm{D}_1\psi \times \mathrm{D}_2\psi\big\|}\,, \quad \mathrm{d}\sigma = \big\|\mathrm{D}_1\psi \times \mathrm{D}_2\psi\big\|\,\mathrm{d}u\,.$$

Après un calcul vu dans le paragraphe 6.3.1, on a

$$\mathrm{D}_1\psi \times \mathrm{D}_2\psi = (a_1, a_2, a_3)$$

avec

$$a_1 = \frac{\partial(\psi_2, \psi_3)}{\partial(u_1, u_2)}\,, \quad a_2 = \frac{\partial(\psi_3, \psi_1)}{\partial(u_1, u_2)}\,, \quad a_3 = \frac{\partial(\psi_1, \psi_2)}{\partial(u_1, u_2)}$$

$$\psi = (\psi_1, \psi_2, \psi_3)\,,$$

ce que l'on écrit symboliquement comme

$$a_1 = \frac{\partial(x_2, x_3)}{\partial(u_1, u_2)}\,, \quad a_2 = \frac{\partial(x_3, x_1)}{\partial(u_1, u_2)}\,, \quad a_3 = \frac{\partial(x_1, x_2)}{\partial(u_1, u_2)}$$

étant sous-entendu que $x_1 = \psi_1$, $x_2 = \psi_2$, $x_3 = \psi_3$. De ce qui précède, l'on voit la cohérence d'écriture symbolique suivante :

$$(\cos\alpha_1)\,\mathrm{d}\sigma = \frac{\partial(x_2, x_3)}{\partial(u_1, u_2)}\,\mathrm{d}u_1\,\mathrm{d}u_2 = \mathrm{d}x_2 \wedge \mathrm{d}x_3$$

$$(\cos\alpha_2)\,\mathrm{d}\sigma = \frac{\partial(x_3, x_1)}{\partial(u_1, u_2)}\,\mathrm{d}u_1\,\mathrm{d}u_2 = \mathrm{d}x_3 \wedge \mathrm{d}x_1$$

$$(\cos\alpha_3)\,\mathrm{d}\sigma = \frac{\partial(x_1, x_2)}{\partial(u_1, u_2)}\,\mathrm{d}u_1\,\mathrm{d}u_2 = \mathrm{d}x_1 \wedge \mathrm{d}x_2\,.$$

Pour l'instant, ces formules sont à considérer comme écriture alternative pour les trois composantes de $\nu\,\mathrm{d}\sigma$. On verra l'efficacité de cette écriture dans les calculs avec les formes différentielles dans le chapitre 7 ; pour l'instant, notons que, l'expression $\nu\,\mathrm{d}\sigma$ étant indépendante (au signe près) du paramétrage (ψ, U) choisi (cf. § 6.3.5), les symboles $\mathrm{d}x_2 \wedge \mathrm{d}x_2$, etc. ont un sens précis (au signe près) sans aucune référence au paramétrage ; le signe lui-même dépend de l'orientation fixée.

Les remarques précédentes montrent que les écritures suivantes pour le flux d'un champ $F = (F_1, F_2, F_3)$ à travers S sont justifiées :

$$\int_S F \cdot \nu \, d\sigma = \int_S (F_1 \cos\alpha_1 + F_2 \cos\alpha_2 + F_3 \cos\alpha_3) \, d\sigma$$
$$= \int_S \{F_1 \, dx_2 \wedge dx_3 + F_2 \, dx_3 \wedge dx_1 + F_3 \, dx_1 \wedge dx_2\} ;$$

notons que l'on écrit $dx_2 \wedge dx_3$ et non $dx_3 \wedge dx_2$; le symbolisme adopté donnera

$$dx_3 \wedge dx_2 = -dx_2 \wedge dx_3$$

car

$$\frac{\partial(x_3, x_2)}{\partial(u_1, u_2)} = -\frac{\partial(x_2, x_3)}{\partial(u_1, u_2)} .$$

Pour l'analyse vectorielle telle qu'elle est développée dans ce chapitre, toutes ces écritures alternatives sont superflues ; nous les avons données en vue de préparer le terrain pour une meilleure compréhension physique du formalisme des formes différentielles donné dans le chapitre suivant. Notons que, dans la littérature technique, l'on trouve aussi l'écriture $d\vec{\sigma}$ pour $\nu \, d\sigma$ de sorte que le flux $\int_S F \cdot \nu \, d\sigma$ s'écrit $\int_S F \cdot d\vec{\sigma}$.

6.3.7 Intégrales de surface : définition globale

Un ensemble non vide $X \subset \mathbb{R}^3$ s'appelle une ***surface régulière*** dans $\mathbb{R}^3$ (de classe C^p, $p \geqslant 1$) si pour tout point $x \in X$, il existe un voisinage ouvert V de x tel que $V \cap X$ est un morceau de surface régulière (de classe C^p) ; une surface régulière (brièvement : une surface) dans $\mathbb{R}^3$ est donc un cas particulier d'une sous-variété différentielle de dimension k dans $\mathbb{R}^N$ (avec $k = 2$, $N = 3$) définie dans le chapitre 4. Le chapitre suivant traitera la théorie générale de l'intégration sur les sous-variétés de dimension k dans $\mathbb{R}^N$; dans ce paragraphe nous donnons quelques indications pour le cas d'une surface dans $\mathbb{R}^3$.

Evidemment, un morceau de surface dans $\mathbb{R}^3$ est un cas très particulier d'une surface dans $\mathbb{R}^3$, cette dernière étant, en quelque sorte, un recollement approprié de morceaux de surfaces. Dans le paragraphe 6.3.3, nous avons défini les intégrales sur un morceau de surface ; les surfaces X dans $\mathbb{R}^3$ que nous rencontrerons dans ce chapitre seront toutes de la forme suivante :

$$X = S_1 \cup \cdots \cup S_n \cup E \tag{6.22}$$

où les S_i sont des morceaux de surfaces et E une réunion d'un nombre fini de points et de courbes (comme définies dans § 6.1.17), les ensembles $S_1, \ldots, S_n, E$ étant deux à deux disjoints. On posera alors, pour $f : X \to \mathbb{R}$,

$$\int_X f \, d\sigma = \sum_{i=1}^{n} \int_{S_i} f \, d\sigma \tag{6.23}$$

en supposant que f est intégrable dans chaque S_i (cf. § 6.3.5) ; on dira alors que f est ***intégrable*** dans X. On peut démontrer que la définition de $\int_X f\,d\sigma$ ne dépend pas de la subdivision de X en morceaux de surfaces S_i ; toute autre subdivision en morceaux de surfaces donnera le même résultat ; ce fait sera établi dans un contexte plus général dans le chapitre suivant. Il est clair suivant cette définition que nous considérons les points et les courbes comme des ensembles négligeables pour la « mesure superficielle » σ ; ceci sera justifié ultérieurement en détail.

Notons qu'un ensemble $X \subset \mathbb{R}^3$ peut être représentable comme dans (6.22) sans que X soit nécessairement une surface (régulière ou non, § 6.3.11) ; même si X n'est pas une surface régulière, nous définirons $\int_X f\,d\sigma$ par la formule (6.23) et nous appellerons cette intégrale « l'intégrale de surface de f sur X ».

Notons encore qu'une surface régulière, selon notre définition, n'est pas nécessairement un ensemble connexe ; par contre, un morceau de surface est un ensemble connexe car c'est l'image d'un domaine de $\mathbb{R}^2$ sous une application continue.

6.3.8 Flux à travers une surface : définition globale

Soit X une surface dans $\mathbb{R}^3$ comme définie dans le paragraphe 6.3.7. Chaque point $x \in X$ étant dans un morceau de surface (contenu dans X), on peut définir sans ambiguïté l'espace tangent T_x et les deux normales unitaires à x suivant la discussion du paragraphe 6.3.4. La surface X s'appelle ***orientable*** si l'on peut choisir une normale unitaire $\nu(x)$ en chaque point $x \in X$ telle que $x \mapsto \nu(x)$ est continue dans X ; on dira alors que $x \mapsto \nu(x)$ est une ***orientation*** pour X. Il est clair que si $x \mapsto \nu(x)$ est une orientation pour X, $x \mapsto -\nu(x)$ est aussi une orientation pour X, appelée l'orientation ***opposée*** à la première. Il est facile de démontrer que *si X est une surface orientable et connexe alors X possède exactement deux orientations, l'une opposée à l'autre.* Nous omettons cette démonstration facile car elle est contenue dans la théorie générale d'orientabilité des sous-variétés de dimension quelconque donnée dans le chapitre suivant.

Contrairement aux courbes qui sont orientées de par leurs définitions, il existe des surfaces dans $\mathbb{R}^3$ qui ne sont pas orientables ; l'exemple le plus simple visuellement est le ***ruban de Möbius***. Cette surface est formée en collant les deux extrémités d'un long ruban rectangulaire après l'avoir tordu de telle manière que le côté $B'A'$ se pose sur le côté AB avec B' sur A et A' sur B ; une telle surface possède « un seul côté » dans le sens qu'en promenant une flèche (la normale unitaire) le long de la courbe simple fermée CC', on obtiendra que la flèche pointe à la direction opposée à son retour au point de départ. Ce fait physique est la preuve de l'impossibilité d'un choix continu de la normale unitaire sur toute la surface du ruban ; un développement analytique général de ce qui précède est renvoyé au chapitre suivant (voir aussi Ex. 19, § 6.3.12 pour le ruban de Möbius).

Soit, maintenant, X une surface orientable dans $\mathbb{R}^3$, orientée par le choix continu d'une normale unitaire $x \mapsto \nu(x)$, $x \in X$; si $F : X \to \mathbb{R}^3$ est un champ vectoriel dans X, le ***flux*** de F à travers X orientée par ν est défini par

$$\int_X (F \cdot \nu)\, d\sigma$$

dans la mesure où la fonction $|F \cdot \nu|$ est intégrable dans X.

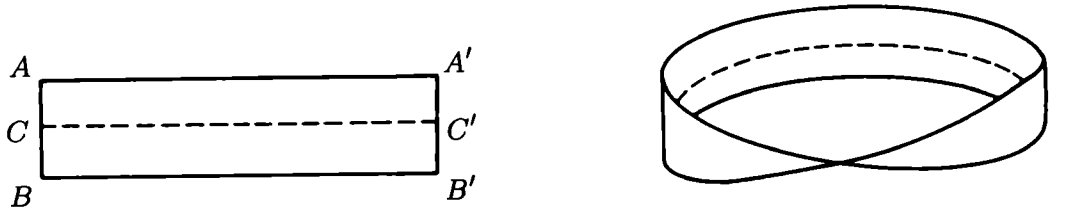


Fig. 6.11

6.3.9 Exemples de surfaces orientables

On sait déjà qu'un morceau de surface régulière dans $\mathbb{R}^3$ est orientable ; on exprime ce fait en disant qu'une surface régulière quelconque dans $\mathbb{R}^3$ est ***localement orientable***.

Soit $f : A \to \mathbb{R}$ une fonction réelle de classe C^1 définie dans une partie ouverte A de $\mathbb{R}^3$; on sait que si c est une valeur régulière de f (ce qui veut dire que $\nabla f(x) \neq 0$ pour tout x tel que $f(x) = c$, cf. § 4.4.7), alors $X = f^{-1}(c)$ *est une surface régulière ; c'est une surface orientable, une orientation étant donnée par*

$$\nu(x) = \frac{\nabla f(x)}{\|\nabla f(x)\|}, \quad x \in X .$$

En effet, si $a \in X$, $a = (a_1, a_2, a_3)$, on a $f(a) = c$ et $\nabla f(a) \neq 0$; ainsi, $D_j f(a) \neq 0$ pour au moins une valeur de $j = 1, 2, 3$; si, par exemple, $D_3 f(a) \neq 0$, on sait, selon le théorème des fonctions implicites, qu'il existe une boule ouverte $B(a; r)$ telle que

$$B(a; r) \cap X = \{(x_1, x_2, x_3) : x_3 = h(x_1, x_2),\ (x_1, x_2) \in V\}$$

où V est un domaine de $\mathbb{R}^2$ contenant (a_1, a_2), $h : V \to \mathbb{R}$ une fonction de classe C^1 et $h(a_1, a_2) = a_3$; en fait, le morceau de surface $B(a; r) \cap X$ est paramétré par

$$(x_1, x_2) \mapsto \big(x_1,\ x_2,\ h(x_1, x_2)\big), \quad (x_1, x_2) \in V ,$$

d'où l'on voit que l'espace tangent T_a est le sous-espace de $\mathbb{R}^3$ engendré par

$$\big(1,\ 0,\ D_1 h(a_1, a_2)\big), \quad \big(0,\ 1,\ D_2 h(a_1, a_2)\big) ;$$

puisque

$$f\big(x_1, x_2, h(x_1, x_2)\big) - c = 0\,, \quad (x_1, x_2 \in V)\,,$$

on obtient que

$$\mathrm{D}_1 h = \frac{-\mathrm{D}_1 f}{\mathrm{D}_3 f}\,, \quad \mathrm{D}_2 h = \frac{-\mathrm{D}_2 f}{\mathrm{D}_3 f}\,,$$

d'où l'on voit que ∇f est orthogonal à T_a ; on obtient ν en normalisant ∇f.

On peut démontrer que *toute surface régulière de classe C^1 orientable contenue dans $\mathbb{R}^3$ est de la forme $f^{-1}(c)$ pour une valeur régulière c d'une fonction f de classe C^1*. En plus, *une surface régulière compacte de classe C^1 dans $\mathbb{R}^3$ est orientable*. Les démonstrations de ces deux affirmations ne seront pas données ici ; voici deux références à ce sujet : H. Samelson, *Proceedings of the American Mathematical Society*, vol. 22 (1969), pp. 301-302 ; E. Lima, *Expositiones Mathematicae*, vol. 5 (1987), pp. 283-286.

Si la surface X est donnée comme $f^{-1}(c)$, nous dirons que X est définie par une ***équation implicite*** $f(x_1, x_2, x_3) = c$. On a déjà vu plusieurs exemples de telles surfaces dans l'exercice 1 (§ 4.4.10).

Les surfaces compactes et connexes dans $\mathbb{R}^3$ s'appellent souvent les surfaces ***closes*** ou ***fermées*** ; nous éviterons cette terminologie qui a néanmoins l'avantage de communiquer l'idée intuitive juste qu'une telle surface renferme un domaine. Plus précisément, on a le théorème suivant : *si X est une surface régulière compacte connexe dans $\mathbb{R}^3$, alors $\mathbb{R}^3$ est la réunion de trois ensembles deux à deux disjoints : $\mathbb{R}^3 = \Omega \cup X \cup \Omega'$ où Ω est un domaine borné et Ω' est un domaine non borné avec $X = \mathrm{Fr}\,\Omega = \mathrm{Fr}\,\Omega'$*. On dira que Ω est le ***domaine intérieur*** défini par X et que Ω' est le ***domaine extérieur*** défini par X. On voit immédiatement que le théorème donné est une généralisation à $\mathbb{R}^3$ du théorème de Jordan dans $\mathbb{R}^2$ (§ 6.2.3) ; ces théorèmes sont des cas particuliers d'un théorème général de Brouwer concernant la séparation de $\mathbb{R}^N$ par des hypersurfaces compactes connexes, théorème démontré dans les livres de topologie algébrique (voir, par exemple, un exposé élémentaire dans Hurewicz et Wallman [22] ou Newman [31]). On peut alors donner un sens simple à la phrase « normale unitaire dirigée vers l'extérieur » en relation avec une surface X compacte et connexe ; c'est une normale unitaire $\nu(x)$ à $x \in X$ telle que $x + \varepsilon\nu(x)$ est dans Ω' pour $\varepsilon > 0$ suffisamment petit. Lorsque une surface compacte connexe X est orientée par la normale unitaire dirigée vers l'extérieur, nous dirons aussi de façon concise que X est orientée ***positivement*** ; avec l'orientation opposée, on dira que X est orientée ***négativement***. Cette terminologie sera utilisée même pour une surface compacte connexe non régulière (cf. § 6.3.11) pourvu que X soit de la forme (6.22) où alors les normales unitaires seront définies seulement sur les morceaux S_i.

Nous parlerons aussi du flux de F à travers X lorsque X n'est pas nécessairement une surface régulière orientable mais est représentable comme dans (6.22), après avoir fixé une orientation pour chacun des morceaux de surfaces S_i qui figurent dans la définition de X et en négligeant essentiellement la définition de ν sur l'ensemble exceptionnel E. Souvent, les orientations des S_i

seront indiquées en décrivant la direction des normales unitaires par rapport à un domaine dont la frontière contient X, comme c'est le cas ci-dessus lorsque X est une surface régulière compacte dans $\mathbb{R}^3$. Lorsque X est donnée par (6.22) et qu'une orientation est fixée pour chacune des S_i, on définira le flux d'un champ vectoriel F à travers S par la formule de type (6.23) :

$$\int_S F \cdot \nu \, d\sigma = \sum_{i=1}^{n} \int_{S_i} F \cdot \nu \, d\sigma .$$

6.3.10 Remarques géométriques sur la définition de l'aire d'une surface

On peut se demander s'il est possible de donner une définition plus géométrique de l'aire d'une surface comme on l'a fait pour la longueur d'une courbe (§ 6.1.7) par une approximation polygonale. Ceci est possible mais l'exemple simple suivant (dû à H. A. Schwarz) montre la nécessité d'une considération attentive du problème ; nous présentons cet exemple géométriquement.

Soit S la surface cylindrique circulaire de rayon 1 et de hauteur 1 ; on peut décrire S analytiquement comme les points $(x, y, z) \in \mathbb{R}^3$ tels que $x^2 + y^2 = 1$ avec $0 < z < 1$. Il est bien connu que l'aire de S est égale à 2π et on obtient ce résultat facilement de notre définition du pararaphe 6.3.3 (cf. Ex. 1(iv), § 6.3.12). Subdivisons le cercle de base de S ($z = 0$) en n arcs égaux et la hauteur de S en m intervalles égaux et considérons les cercles de S correspondants à $z = k/m$, $k = 0, 1, \ldots, m-1$; subdivisons chacun de ces m cercles en n arcs égaux de telle sorte que les points de subdivision du cercle avec $z = k/m$ projettent verticalement aux milieux des arcs de subdivision du cercle avec $z = (k-1)/m$. Soit $P_{m,n}$ le polyèdre inscrit dans S dont les faces sont les $2mn$ triangles isocèles suivants ; écrivons :

$$\xi_j^k = \left(\cos \frac{2j\pi}{n} , \sin \frac{2j\pi}{n} , \frac{k}{m} \right)$$

$$\eta_j^k = \left(\cos \frac{(2j+1)\pi}{n} , \sin \frac{(2j+1)\pi}{n} , \frac{k}{m} \right)$$

avec $j = 0, 1, \ldots, n$, $k = 0, 1, \ldots, m$; soit A_j^k le triangle de sommets $\{\xi_j^k, \xi_{j+1}^k, \eta_j^{k+1}\}$ et B_j^k celui de sommets $\{\eta_j^{k+1}, \eta_{j+1}^{k+1}, \xi_j^k\}$ avec $0 \leqslant j \leqslant n-1$ et $0 \leqslant k \leqslant m-1$. On vérifie aisément que ces $2mn$ triangles ont la même aire égale à celle du triangle A_0^0 qui a les sommets

$$\xi_0^0 = (1,0,0) , \quad \xi_1^0 = \left(\cos \frac{2\pi}{n} , \sin \frac{2\pi}{n} , 0 \right) , \quad \eta_0^1 = \left(\cos \frac{\pi}{n} , \sin \frac{\pi}{n} , \frac{1}{m} \right) ;$$

la base de A_0^0 est de longueur $b = \|\xi_0^0 - \xi_1^0\| = 2\sin(\pi/n)$ et les deux côtés égaux sont de longueur

$$a = \|\xi_0^0 - \eta_0^1\| = \left(4 \sin^2 \frac{\pi}{2n} + \frac{1}{m^2} \right)^{1/2} ;$$

donc l'aire A_0^0 vaut (formule de Heron) :

$$\frac{b}{4}\sqrt{4a^2-b^2} = \frac{1}{m}\sin\frac{\pi}{n}\left\{1+4m^2\left(\sin\frac{\pi}{2n}\right)^4\right\}^{1/2}$$

d'où l'on obtient l'aire α_{mn} de P_{mn}

$$\alpha_{mn} = 2mn\cdot \text{aire}(A_0^0) = 2n\sin\frac{\pi}{n}\left\{1+4m^2\left(\sin\frac{\pi}{2n}\right)^4\right\}^{1/2}$$

Si n est fixé mais $m \to \infty$, α_{mn} tend vers ∞ ; si m et n tendent vers ∞ de telle manière que $mn^{-2} \to c \geqslant 0$, alors α_{mn} tend vers

$$2\pi\left\{1+\frac{\pi^4}{4}c^2\right\}^{1/2} ;$$

cette dernière expression peut prendre toutes les valeurs dans $[2\pi, \infty]$ lorsque c varie dans $[0, \infty]$. On constate donc que les aires des polyèdres inscrits dans S ne tendent pas vers une valeur unique. Néanmoins, on voit que la borne inférieure des aires des polyèdres inscrits ci-dessus est 2π, la vraie aire de S ; une définition géométrique d'aire d'une surface L proposée par Lebesgue consiste effectivement à prendre « la borne inférieure des aires des surfaces polyédrales qui tendent uniformément vers L » (Lebesgue 1902 ; cf. ses œuvres scientifiques). On peut démontrer que cette définition aboutit à la formule de l'aire d'un morceau de surface régulière proposée dans le paragraphe 6.3.3.

6.3.11 Sur la définition d'une surface

Ce que nous avons défini comme étant une surface dans $\mathbb{R}^3$ est un sous-ensemble de $\mathbb{R}^3$ qui est localement homéomorphe à une partie ouverte de $\mathbb{R}^2$ et qui possède un plan tangent à chaque point variant continûment lorsque l'on passe d'un point à un autre de la surface. Si l'on considère les sous-ensembles de $\mathbb{R}^3$ qui sont seulement localement homéomorphes à une partie ouverte de $\mathbb{R}^2$, on obtient les ***sufaces topologiques*** dans $\mathbb{R}^3$ non nécessairement régulières. Ainsi, la frontière d'un pavé borné dans $\mathbb{R}^3$ est une surface topologique non régulière car elle possède des arêtes et des sommets anguleux où un plan tangent ne peut pas exister ; néanmoins, c'est une surface qui n'est pas trop compliquée dans le sens où elle est la réunion d'un nombre fini (6) de morceaux de surfaces deux à deux disjoints et un ensemble E formé de 12 arêtes et de 8 sommets, donc contenu dans une réunion finie de courbes ; du point de vue de l'aire superficielle, E est négligeable, de sorte que l'on peut intégrer, sans difficulté, sur la surface formée par la frontière d'un pavé borné. Nous appellerons ce type de surface, une surface régulière par morceaux ; plus précisément, un ensemble $S \subset \mathbb{R}^3$ sera appelé une ***surface régulière par morceaux*** si S est une surface topologique et si

$$S = S_1 \cup \cdots \cup S_n \cup E$$

où chaque S_i est une surface régulière (éventuellement un morceau de surface régulière), les ensembles $S_1, \ldots, S_n$, E étant deux à deux disjoints et E une réunion finie de points et de courbes. Une intégrale de surface sur S sera *par définition* la somme des n intégrales de surface sur $S_1, \ldots, S_n$ respectivement. Les points de S qui sont dans un S_i seront appelés les ***points réguliers*** de S.

Notons encore que ni S ni les S_i sont nécessairement des ensembles connexes.

Il faut comprendre que la plupart des surfaces dans $\mathbb{R}^3$ ne sont pas homéomorphes à une partie ouverte de $\mathbb{R}^2$, elles sont seulement localement homéomorphes aux parties ouvertes de $\mathbb{R}^2$. Ainsi, une sphère S dans $\mathbb{R}^3$ ne peut pas être homéomorphe à une partie ouverte U de $\mathbb{R}^2$ car S est compact et U (non vide) ne l'est pas ; en fait, S ne peut pas être homéomorphe à aucun sous-ensemble de $\mathbb{R}^2$; la preuve de cette affirmation est un exercice simple de topologie. Mais on peut avoir des surfaces intéressantes dans $\mathbb{R}^3$ qui sont homéomorphes à certains sous-ensembles ouverts de $\mathbb{R}^2$; par exemple, un cylindre circulaire dans $\mathbb{R}^3$ est homéomorphe à $\mathbb{R}^2 \setminus \{(0,0)\}$ (cf. Ex. 18, § 6.3.12).

6.3.12 Exercices

1 Faire les vérifications indiquées ci-dessous (les points de $\mathbb{R}^3$ étant notés (x,y,z)) :

(i) soit $S = \{(x,y,z) : ax + by + cz = d\}$ (un plan) où a, b, c, d sont réels avec $c \neq 0$; S est un morceau de surface d'équation cartésienne $z = (d - ax - by)/c$, $(x,y) \in \mathbb{R}^2$; si $f : S \to \mathbb{R}$,

$$\int_S f \, d\sigma = k \int_{\mathbb{R}^2} f\left(x, y, \frac{d - ax - by}{c}\right) dx\, dy$$

où $k = (a^2 + b^2 + c^2)^{1/2}/|c|$; si $A \subset S$, on a $\sigma(A) = k\lambda_2(\tilde{A})$ où $\tilde{A}$ est l'image de A par la projection orthogonale sur le plan (x,y) ; en écrivant une normale unitaire à un point de S comme $\nu = (\cos\alpha, \cos\beta, \cos\gamma)$, on obtient $k = |\sec\gamma|$;

(ii) soit $S = \{(x,y,z) : x^2 + y^2 + z^2 = a^2\}$ (une sphère) où $a > 0$; on a $S = S_+ \cup S_- \cup E$ où E est la partie de S avec $z = 0$ (E est « l'équateur ») et S_+, S_- sont les parties de S avec $z > 0$, $z < 0$ respectivement ; S_+ est un morceau de surface (hémisphère nord) d'équation cartésienne $z = \sqrt{a^2 - x^2 - y^2}$, (x,y) dans $W = \{(x,y) : x^2 + y^2 < 1\}$; S_- est un morceau de surface (hémisphère sud) d'équation cartésienne $z = -\sqrt{a^2 - x^2 - y^2}$, $(x,y) \in W$; on a $d\sigma = a/\sqrt{a^2 - x^2 - y^2}\, dx\, dy$ pour S_+ et S_- et $\sigma(S_+) = \sigma(S_-) = 2\pi a^2$ avec $\sigma(S) = 4\pi a^2$; une normale unitaire au point $(x,y,z) \in S$ est un vecteur proportionnel à (x,y,z) ;

(iii) soit S comme dans (ii) ; si E est la partie de S où $x \geqslant 0$ et $y = 0$ (E est « le méridien de Greenwich ») alors $S \setminus E$ est un morceau de surface paramétré par

$$\psi(\theta, \varphi) = a(\sin\theta\cos\varphi,\ \sin\theta\sin\varphi,\ \cos\theta)\,, \quad 0 < \theta < \pi\,,\ 0 < \varphi < 2\pi\,,$$

(paramétrage sphérique) avec $\mathrm{d}\sigma = a^2 \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi$; si A est la partie de S définie par $\psi^{-1}(A) = \,]\,\theta_1\,,\,\theta_2\,[\,\times\,]\,\varphi_1\,,\,\varphi_2\,[$, alors

$$\sigma(A) = a^2(\varphi_2 - \varphi_1)(\cos\theta_1 - \cos\theta_2)$$

d'où $\sigma(A_\theta) = 2\pi a^2(1 - \cos\theta)$ où A_θ est la partie de S entre le « pôle nord » $(0, 0, a)$ et un cercle de « co-latitude » θ (A_θ s'appelle une calotte) ;

(iv) soit $S = \{(x, y, z) : x^2 + y^2 = a^2\}$ (un cylindre circulaire) où $a > 0$; si E est la partie de S où $x = a$ (E est une droite « verticale »), alors $S \setminus E$ est un morceau de surface paramétré par

$$\psi(\theta, z) = (a\cos\theta,\ a\sin\theta,\ z)\,, \quad 0 < \theta < 2\pi\,,\ z \in \mathbb{R},$$

avec $\mathrm{d}\sigma = a\,\mathrm{d}\theta\,\mathrm{d}z$ et $\sigma(A_h) = 2\pi a h$ où A_h est une partie de S entre les plans $z = z_0$ et $z = z_0 + h$ ($h > 0$; A_h est un cylindre circulaire de rayon de base a et de hauteur h) ; si S_+ est la partie de S où $y > 0$ et S_- celle où $y < 0$, alors les morceaux de surfaces $S_\pm$ sont donnés par les équations cartésiennes $y = \pm\sqrt{a^2 - x^2}$, (x, z) dans $W = \{(x, z) : -a < x < a\,,\ z \in \mathbb{R}\}$ avec $\mathrm{d}\sigma = a/\sqrt{a^2 - x^2}\,\mathrm{d}x\,\mathrm{d}z$ et $S = S_+ \cup S_- \cup E$ où E est la partie de S où $y = 0$ (E est la réunion de deux droites verticales) ; en fait, S tout entier est un morceau de surface (voir exercice 18) ;

(v) soit $S = \{(x, y, z) : x^2 + y^2 - k^2 = 0\}$ (un cône circulaire) où $k > 0$ ($k = \mathrm{tg}\,\alpha$, $0 < \alpha < \pi/2$, α étant l'angle au « sommet » $(0, 0, 0)$ entre l'axe des z, appelé l'axe de symétrie du cône, et une « droite génératrice » du cône) ; si E est la droite génératrice $(tk, 0, t)$, $t \in \mathbb{R}$, alors $S \setminus E$ est un morceau de surface paramétré par

$$\psi(z, \theta) = \big(kz\cos\theta,\ kz\sin\theta,\ z\big)\,, \quad z \in \mathbb{R}^*\,,\ 0 < \theta < 2\pi\,,$$

avec $\mathrm{d}\sigma = |z|\,k\sqrt{1 + k^2}\,\mathrm{d}z\,\mathrm{d}\theta$; si A est la partie de S où $0 \leqslant z \leqslant h$, on aura $\sigma(A) = \pi k\sqrt{1 + k^2}\,h^2$, ce qui vaut πRL en posant $R = h\,\mathrm{tg}\,\alpha$ et $L = h\sec\alpha$ (R étant le rayon du cercle de base du cône A de hauteur h et de longueur latérale L) ; calculer une normale unitaire ν à un point de $S \setminus E$ par la formule analytique en remarquant qu'elle correspond à notre intuition géométrique ; si S_+ est la partie de S où $z > 0$ et S_- celle de $z < 0$, leurs équations cartésiennes respectives sont $z = \pm(1/k)\sqrt{x^2 + y^2}$, $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ avec $\mathrm{d}\sigma = \sqrt{k^2 + 1}/k\,\mathrm{d}x\,\mathrm{d}y$; revérifier la formule pour $\sigma(A)$ (A étant le cône ci-dessus) en utilisant ce $\mathrm{d}\sigma$.

2 Déterminer les aires suivantes (les points de $\mathbb{R}^3$ étant notés (x,y,z)) :

(i) l'aire de la partie du plan d'équation $6x + 3y + 2z = 12$ contenue dans l'octant positif ;

(ii) l'aire de la partie bornée du cône d'équation $x^2 + y^2 = z^2$ interceptée par le cylindre d'équation $x^2 + y^2 = 2ax$ $(a > 0)$;

(iii) l'aire de la partie bornée du paraboloïde d'équation $2z = x^2+y^2$ interceptée par le cylindre donné par $x^2 + y^2 = 1$;

(iv) l'aire de la partie du cylindre d'équation $x^2 + y^2 = ax$ contenue dans le domaine intérieur défini par la sphère donnée par $x^2+y^2+z^2 = a^2$ $(a > 0)$;

(v) l'aire de la partie de la sphère d'équation $x^2 + y^2 + z^2 = a^2$ interceptée par le cylindre donné par $x^2 - ax + y^2 = 0$ $(a > 0)$;

(vi) l'aire de la partie bornée du cylindre d'équation $x^2 + y^2 = a^2$ interceptée par le cylindre donné par $y^2 + z^2 = a^2$.

3 Calculer les intégrales suivantes (les points de $\mathbb{R}^3$ étant notés (x,y,z)) :

(i) $\int_S xyz \, d\sigma$, S étant la partie du plan d'équation $x + y + z = 1$ contenue dans l'octant positif ;

(ii) $\int_S (x^2+y^2)z \, d\sigma$, S étant la partie de la sphère d'équation $x^2+y^2+z^2 = 4$ où $z \geqslant 1$;

(iii) $\int_S \sqrt{x^2+y^2} \, d\sigma$, S étant le morceau de surface hélicoïdale donné par $x = r\cos\theta$, $y = r\sin\theta$, $z = \theta$, $0 < r < 1$, $0 < \theta < 2\pi$;

(iv) $\int_S z^2/\sqrt{1+x^2+y^2} \, d\sigma$, S étant la partie bornée du paraboloïde hyperbolique d'équation $z = xy$ interceptée par le cylindre d'équation $x^2 + y^2 = R^2$;

(v) $\int_S x^2/\sqrt{1+4z} \, d\sigma$, S étant la surface donnée par $z = x^2+y^2$, $x^2+y^2 < 1$.

4 Calculer le flux du champ vectoriel F à travers S dans les cas suivants (les points de $\mathbb{R}^3$ étant notés (x,y,z)) :

(i) $F(x,y,z) = (xy, xz, z^3+x^2+y^2)$, S étant la partie du cylindre d'équation $x^2 + y^2 = 1$ où $0 < z < 1$, orientée par la normale dirigée vers l'extérieur (du domaine déterminé par $x^2 + y^2 < 1$) ;

(ii) $F(x,y,z) = (xy, yz, xz)$, S étant la pyramide formée des plans d'équations $x = 0$, $y = 0$, $z = 0$ et $x+y+z = 1$ respectivement, orientée par la normale dirigée vers l'extérieur de la pyramide ;

(iii) $F(x,y,z) = (x,y,z)$, S étant la partie où $z > 0$ d'un cône circulaire de sommet $(h,0,0)$, $h \geqslant 0$, ayant pour axe de symmétrie l'axe des z et interceptant le plan (x,y) dans un cercle de centre $(0,0,0)$ et de rayon $R > 0$; S est orientée positivement par rapport au volume limité par S et le plan (x,y) (remarquer que S satisfait $z = h - (h/R)\sqrt{x^2+y^2}$) ;

(iv) $F(x,y,z) = (xz, yz, 1)$, S étant l'hémisphère nord $(z > 0)$ de la sphère centrée en $(0,0,0)$ et rayon $R > 0$, orienté positivement par rapport au volume limité par la sphère.

5 Calculer les intégrales suivantes (les points de $\mathbb{R}^3$ étant notés (x,y,z)) en tenant compte des notations du paragraphe 6.3.6 :

(i) $\int_S x\,dy\wedge dz$, S étant la partie du plan $x+2y+z=1$ dans l'octant positif orientée par la normale $(1,2,1)$;

(ii) $\int_S x^2y^2z\,dx\wedge dy$, S étant l'hémisphère sud ($z<0$) de la sphère d'équa-tion $x^2+y^2+z^2=R^2$ ($R>0$), orienté positivement en tant que partie de la sphère ;

(iii) $\int_S z\,dz\wedge dx$, S étant la partie du cylindre d'équation $x^2+y^2=a^2$ entre $z=0$ et $z=1$, orientée par la normale dirigée vers l'extérieur du volume limité par le cylindre ;

(iv) $\int_S z^2\,dx\wedge dz$, S étant la partie du cône d'équation $x^2+y^2-z^2=0$ entre $z=0$ et $z=1$, orientée par la normale dirigée vers l'extérieur du volume intérieur du cône.

6 Soit S un morceau de surface régulière donné par une équation cartésienne $z=f(x,y)$, $(x,y)\in U$, U une partie ouverte de $\mathbb{R}^2$, orienté par une normale unitaire ν telle que $\nu\cdot(0,0,1)>0$. Si $F=(F_0,F_0,F_3)$ est un champ vectoriel défini sur S, vérifier que le flux de F à travers S est égal à

$$\iint_U F_3\big(x,\,y,\,f(x,y)\big)\,dx\,dy$$

pourvu que cette intégrale existe.

7 Soit S la sphère centrée en $0\in\mathbb{R}^3$ de rayon $R>0$, orientée positivement ; soit $F(x)=f\big(\|x\|\big)x$, $x\neq 0$, un champ vectoriel (dit *central*) où $f:\,]0\,,\,\infty[\,\to\mathbb{R}$. Si $A\subset S$, vérifier que le flux de F à travers A est égal à $Rf(R)\sigma(A)$.

8 Soient S la sphère centrée en $0\in\mathbb{R}^3$ de rayon $R>0$ et T le cylindre fini défini par $x^2+y^2=R^2$, $-R<z<R$; soient α, β les paramétrages de S moins un méridien et de T moins un segment, donnés respectivement par

$$\alpha(\theta,\varphi)=R(\sin\theta\cos\varphi,\,\sin\theta\sin\varphi,\,\cos\theta)$$
$$\beta(\theta,\varphi)=R(\cos\varphi,\,\sin\varphi,\,\cos\theta)$$

où $0<\theta<\pi$, $0<\varphi<2\pi$. Vérifier que l'application $f:S\to T$ donnée par $f\big(\alpha(\theta,\varphi)\big)=\beta(\theta,\varphi)$ est bien définie et que si $A\subset S$, alors $\sigma(A)=\sigma\big(f(A)\big)$; en particulier, $\sigma(S)=\sigma(T)$.

9 Soit C une courbe régulière simple dans $\mathbb{R}^2$ définie par le paramétrage $t\mapsto\big(\alpha(t),\beta(t)\big)$, $a\leqslant t\leqslant b$, éventuellement une courbe de Jordan. Vérifier que $\psi(t,z)=\big(\alpha(t),\beta(t),z\big)$, $a<t<b$, $z_1<z<z_2$, définit le paramétrage d'un morceau de surface régulière dans $\mathbb{R}^3$ avec $d\sigma=\sqrt{\alpha'^2+\beta'^2}\,dt\,dz$; en déduire que l'aire de la surface cylindrique associée à ce morceau de surface est $L(z_2-z_1)$ où L est la longueur de la courbe C.

10 Soit C comme dans l'exercice 9 ; posons

$$\psi(t,u) = \big(u\alpha(t), u\beta(t), u\big)\,.$$

Vérifier que pour $a < t < b$, $0 < u < \infty$, ψ définit un morceau de surface régulière avec

$$\mathrm{d}\sigma = u\left\{\alpha'^2 + \beta'^2 + \big(\alpha'\beta - \beta'\alpha\big)^2\right\}^{1/2} \mathrm{d}t\,\mathrm{d}u\,.$$

(Ce morceau de surface est une partie du cône défini par $h(x,y,z) = 0$ où h est une fonction homogène de degré 0 en x, y, z avec $h(x,y,z) = f(x/z\,,\, y/z)$, f étant telle que $f(x,y) = 0$ si $(x,y) \in C$.)

REMARQUE. Lorsque h est un polynôme homogène du second degré, on obtient les cônes de la géométrie élémentaire.

11 Soit C comme dans l'exercice 9, avec $\alpha(t) > 0$, $a < t < b$; posons

$$\psi(t,\theta) = \big(\alpha(t)\cos\theta,\, \alpha(t)\sin\theta,\, \beta(t)\big)\,, \quad a < t < b\,,\ 0 < \theta < 2\pi\,.$$

Vérifier que ψ définit un morceau de surface régulière dans $\mathbb{R}^3$ (ici § 6.3.2 peut être utile) avec $\mathrm{d}\sigma = \alpha(t)\{\alpha'^2(t) + \beta'^2(t)\}^{1/2}\,\mathrm{d}t\,\mathrm{d}\theta$. Si S est l'ensemble dans $\mathbb{R}^3$ (les points de $\mathbb{R}^3$ étant notés (x,y,z)) engendré par la rotation complète de C autour de l'axe des z, C étant considérée comme une partie du plan (x,z), remarquer que S est donné par ce morceau de surface à un ou deux cercles près ; établir que $\sigma(S) = (2\pi\overline{x})L$ où L est la longueur de C et $\overline{x} = (1/L)\int_C x\,\mathrm{d}s$ est la x-composante du barycentre de C dans le plan (x,z). Ceci donne la *règle de Guldin-Pappus* (pour les aires) : *l'aire d'une surface S obtenue par la révolution d'une courbe plane C autour d'un axe de son plan qui ne l'intersecte pas est égale au produit de la longueur L de C et de la longueur du chemin parcouru par le barycentre de C pendant la révolution.* Constater que la discussion reste valable si $\alpha(t) = 0$ pour un nombre fini de points t même si S n'est plus nécessairement une surface régulière.

12 (Notation de l'exercice 11 maintenue) Si l'équation implicite de C est $f(x,z) = 0$ où f est une fonction de classe C^1 définie pour tout $x > 0$ pour laquelle 0 est une valeur régulière, constater que la surface S possède l'équation implicite $g(x,y,z) = 0$ où $g(x,y,z) = f\big(\sqrt{x^2+y^2}\,,\, z\big)$, de sorte que S est une surface régulière dans $\mathbb{R}^3$. Si

$$f(x,z) = (x-b)^2 + z^2 - a^2\,, \quad 0 < a < b\,,$$

S est un tore ayant une équation implicite polynomiale

$$\big(x^2 + y^2 + z^2 + b^2 - a^2\big)^2 - 4b^2(x^2+y^2) = 0\,;$$

vérifier que l'aire de la surface du tore est égale à $(2\pi b)(2\pi a) = 4\pi^2 ab$.

13 Justifier, en utilisant une discussion similaire à celle de l'exercice 11, la formule connue suivante :

$$\sigma(S) = 2\pi \int_a^b f(x)\sqrt{1 + f'^2(x)}\,\mathrm{d}x$$

où S est la surface de révolution engendrée par la rotation autour de l'axe des x de la courbe d'équation $y = f(x)$, $a < x < b$, avec f de classe C^1 et $f(x) > 0$ pour $a < x < b$; remarquer la relation de cette formule avec la règle de Guldin-Pappus donnée dans l'exercice 11.

14 Soit $\psi(r, \theta) = \big(r\cos\theta,\, r\sin\theta,\, h(\theta)\big)$, $0 < r < R$, $0 < \theta < 2\pi$, avec $h : [\,0\,,\, 2\pi] \to \mathbb{R}$ une fonction continue, de classe C^1 dans $]\,0\,,\, 2\pi\,[$. Vérifier que ψ définit un morceau de surface régulière S dans $\mathbb{R}^3$ avec

$$\mathrm{d}\sigma = \big(r^2 + h'^2(\theta)\big)\,\mathrm{d}r\,\mathrm{d}\theta\,.$$

Si $h(\theta) = c\theta$ $(c \neq 0)$, S est une partie d'une surface hélicoïdale.

REMARQUE. En général, S s'appelle un *conoïde* ; elle est engendrée par un segment mobile qui s'appuie à l'axe des z et qui est toujours parallèle au plan $z = 0$; c'est un exemple d'une *surface réglée* engendrée par le mouvement approprié des segments.

15 Par définition, le barycentre d'une surface $S \subset \mathbb{R}^3$ est $(\overline{x}_1, \overline{x}_2, \overline{x}_3)$ où

$$\overline{x}_i = \frac{1}{\sigma(S)} \int_S x_i\,\mathrm{d}\sigma\,, \quad i = 1, 2, 3\,;$$

démontrer que le barycentre de S, l'hémisphère nord $(x_3 > 0)$ de la sphère d'équation $x_1^2 + x_2^2 + x_3^2 = a^2$ $(a > 0)$, est donné par $(0, 0, a/2)$.

16 Calculer

$$I = \int_S \rho(y^2 + z^2)\,\mathrm{d}\sigma$$

où S est l'hémisphère nord $(z > 0)$ de la sphère d'équation $x^2 + y^2 + z^2 = a^2$ $(a > 0)$ et $\rho > 0$ est une constante (I s'appelle le *moment d'inertie* de S par rapport à l'axe des x, la « densité » de S étant la constante ρ) ; vérifier que $I = (2/3)\,a^2 M$ où M est la masse de S.

17 Calculer

$$I = \int_S \frac{\rho\,\mathrm{d}\sigma}{\big[x^2 + y^2 + (z + a)^2\big]}\,,$$

$\rho > 0$, $a > 0$ étant des constantes et S identique à celui de l'exercice 16 (I peut être interprété comme le potentiel électrostatique au point $(0, 0, -a)$ d'une charge $c = \rho(2\pi a^2)$ uniformément répartie sur l'hémisphère S) ; constater que $I = (c/a)(2 - \sqrt{2})$.

18 Posons $U = \mathbb{R}^2 \setminus \{(0,0)\}$,

$$S_1 = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 = 1\},$$
$$S_2 = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 - x_3^2 = 1\},$$
$$S_3 = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 - x_3^2 = 0,\ x_3 > 0\}$$

où $x = (x_1, x_2, x_3)$. En écrivant $r = (x_1^2 + x_2^2)^{1/2}$ et

$$\psi(x_1, x_2) = \left(\frac{x_1}{r}, \frac{x_2}{r}, \theta(r)\right), \quad (x_1, x_2) \in U$$

où θ est un homéomorphisme de $]0, \infty[$ sur $\mathbb{R}$, vérifier que ψ est un homéomorphisme de U sur S_1 avec

$$\psi^{-1}(x_1, x_2, x_3) = (x_1\theta^{-1}(x_3), x_2\theta^{-1}(x_3)), \quad x \in S_1.$$

En prenant $\theta(t) = \ln t$, $t > 0$, (ψ, U) est même un paramétrage de S_1. Vérifier que

$$f(x_1, x_2, x_3) = (x_1\sqrt{1 + x_3^2}, x_2\sqrt{1 + x_3^2}, x_3)$$

définit un homéomorphisme $f : S_1 \to S_2$ et déterminer $f^{-1} : S_2 \to S_1$. Vérifier que $f(x_1, x_2) = (x_1, x_2, (x_1^2 + x_2^2)^{1/2})$ définit un homéomorphisme $f : U \to S_3$ et déterminer $f^{-1} : S_3 \to U$. Conclure que U, S_1, S_2 et S_3 sont C^1-difféomorphes entre eux.

19 (Ruban de Möbius) Soit

$$\psi(u, v) = \alpha(u) + v\beta(u), \quad u \in \mathbb{R},\ v \in \left]-\frac{1}{2}, \frac{1}{2}\right[$$

où

$$\alpha(u) = (\cos u, \sin u, 0), \quad \beta(u) = \left(\cos\frac{u}{2}\right)\alpha(u) + \left(\sin\frac{u}{2}\right)e_3$$

avec $e_3 = (0, 0, 1)$. Posons

$$M = \psi\left(\mathbb{R} \times \left]-\frac{1}{2}, \frac{1}{2}\right[\right);$$

l'objectif de cet exercice est d'établir que M est une surface non orientable dans $\mathbb{R}^3$ en raisonnant comme suit :

(i) pour tout u, v, $\psi(u + 2\pi, v) = \psi(u, -v)$; pour tout u, $\{\alpha(u), \alpha'(u), e_3\}$ forment une base orthogonale positive pour $\mathbb{R}^3$;

(ii) pour tout $a \in \mathbb{R}$, ψ est injective dans $[a, a + 2\pi[\times [-1/2, 1/2]$;

(iii) au point (u, v),

$$\mathrm{D}_1\psi \times \mathrm{D}_2\psi = \Big(\sin\frac{u}{2}\Big)\Big(1 + v\cos\frac{u}{2}\Big)\alpha(u) + \frac{v}{2}\,\alpha'(u) - \Big(\cos\frac{u}{2}\Big)\Big(1 + v\cos\frac{u}{2}\Big)e_3\,,$$

$$\big\|\mathrm{D}_1\psi \times \mathrm{D}_2\psi\big\|^2 = \frac{v^2}{4} + \Big(1 + v\cos\frac{u}{2}\Big)^2\ ;$$

(iv) Posons

$$S_1 = \psi\Big(\big[0\,,\,2\pi\,\big[\,\times\,\big]-\frac{1}{2}\,,\,\frac{1}{2}\,\big[\Big)\,,\quad S_2 = \psi\Big(\big]\frac{\pi}{2}\,,\,\frac{\pi}{2} + 2\pi\,\Big[\,\times\,\big]-\frac{1}{2}\,,\,\frac{1}{2}\,\big[\Big)\ ;$$

S_1, S_2 sont deux morceaux de surface régulière avec $M = S_1 \cup S_2$, ce qui démontre que M est une surface régulière.

(Pour la propriété homéomorphique de ψ, utiliser les remarques du paragraphe 6.3.2.)

(v) Soit ν le vecteur unitaire défini par la normalisation de $\mathrm{D}_1\psi \times \mathrm{D}_2\psi$; on a

$$\lim_{u\to 0_+} \nu\big(\psi(u,0)\big) = -e_3\,,\qquad \lim_{u\to 2\pi_-} \nu\big(\psi(u,0)\big) = e_3\,.$$

Déduire de (v) que l'on ne peut pas avoir une normale unitaire continue définie partout sur M.

REMARQUE. Imaginons le cercle unité tracé par α dans le plan d'équation $x_3 = 0$ et plaçons une aiguille AB de longueur 1 (soit l'intervalle $]-1/2\,,\,1/2\,[$) avec son milieu au point $\alpha(u)$, le long du vecteur $\beta(u)$; on obtient ainsi le ruban M ; noter que les relations $\psi(0,v) = (1+v)(1,0,0)$, $\psi(2\pi,v) = (1-v)(1,0,0)$ montrent que l'aiguille orientée comme AB à $u = 0$ devient orientée comme BA à $u = 2\pi$.

20 Soit $f : \mathbb{R}^3 \to [\,0\,,\,\infty\,[$ une fonction borélienne ; soit $S(a;r)$ la sphère centrée en a de rayon $r > 0$ et posons $S_r = S(0;r)$. Vérifier les affirmations suivantes :

$$\text{(i)}\ \int_{S_r} f\,\mathrm{d}\sigma = r^2 \int_{S_1} f_r\,\mathrm{d}\sigma \quad \text{où } f_r(x) = f(rx)\ ;$$

$$\text{(ii)}\ \int_{\mathbb{R}^3} f(x)\,\mathrm{d}x = \int_0^\infty \Big(\int_{S_r} f\,\mathrm{d}\sigma\Big)\,\mathrm{d}r = \int_0^\infty r^2 \Big(\int_{S_1} f_r\,\mathrm{d}\sigma\Big)\,\mathrm{d}r\,.$$

Les relations (i) et (ii) restent valables si $f : \mathbb{R}^3 \to \mathbb{R}$ (borélienne) est intégrable (sur S_r pour (i), dans $\mathbb{R}^3$ pour (ii)). Déduire de (ii) que

$$\int_{B(a;R)} f(x)\,\mathrm{d}x = \int_0^R \Big(\int_{S(a;r)} f\,\mathrm{d}\sigma\Big)\,\mathrm{d}r\,.$$

(Pour (i) utiliser les paramétrages sphériques pour S_r et S_1 ; pour (ii) utiliser la transformation sphérique dans $\mathbb{R}^3$.)

6.4 THÉORÈME DE GAUSS DANS $\mathbb{R}^3$

Le théorème de Gauss dans $\mathbb{R}^3$, appelé aussi ***théorème de Gauss-Ostrogradsky*** ou ***théorème de la divergence***, concerne le flux des champs vectoriels dans $\mathbb{R}^3$ à travers les surfaces fermées ; il est une généralisation naturelle du théorème de Green dans $\mathbb{R}^2$ dans sa forme (6.16) où il s'agit du flux à travers une courbe plane fermée.

6.4.1 Domaines réguliers dans $\mathbb{R}^3$

Soit Ω un domaine dans $\mathbb{R}^3$; par définition, Ω est un ensemble ouvert, non vide, connexe ; même si Ω est connexe, $\operatorname{Fr}\Omega$ n'est pas nécessairement connexe et même pour un domaine borné Ω quelconque, $\operatorname{Fr}\Omega$ peut être très compliqué. Nous allons considérer une classe de domaines bornés dans $\mathbb{R}^3$ dont les frontières sont relativement simples ; leur description est analogue à celle des domaines bornés réguliers dans $\mathbb{R}^2$ donnés dans le paragraphe 6.2.4. Appelons un domaine borné $\Omega \subset \mathbb{R}^3$ un domaine ***simple*** si $\operatorname{Fr}\Omega = S$ où S est une surface connexe régulière par morceaux (cf. § 6.3.11) possédant la propriété suivante : si $x \in S$ appartient à une surface régulière $S_i \subset \operatorname{Fr}\Omega$ (x est un point régulier de S), on a un $\varepsilon_0 > 0$ et une normale unitaire $\nu(x)$ au point x de la surface S_i tels que

$$x + \varepsilon\nu(x) \in \operatorname{ext}\Omega\,, \quad x - \varepsilon\nu(x) \in \Omega\,, \quad 0 < \varepsilon < \varepsilon_0\,.$$

Nous dirons que ν est la normale unitaire ***dirigée vers l'extérieur*** et $-\nu$ celle ***dirigée vers l'intérieur*** (***de*** Ω). Si F est un champ vectoriel défini sur S, on définira le ***flux*** de F à travers S (vers l'extérieur) comme

$$\int_S F \cdot \nu \, \mathrm{d}\sigma = \sum_{i=1}^{n} \int_{S_i} F \cdot \nu \, \mathrm{d}\sigma$$

où $S = S_1 \cup \cdots \cup S_n \cup E$, E étant l'ensemble exceptionnel du type indiqué dans le paragraphe 6.3.11.

Un domaine borné simple permet de définir un domaine ***régulier*** dans $\mathbb{R}^3$. Si Ω est un domaine borné dans $\mathbb{R}^3$ tel que

$$\Omega = D_1 \setminus \left(\overline{D}_2 \cup \cdots \cup \overline{D}_n\right)$$

où $D_1, \ldots, D_n$ sont n domaines bornés simples, $\overline{D}_2, \ldots, \overline{D}_n$ étant deux à deux disjoints et contenus dans D_1, alors nous dirons que Ω est un domaine ***régulier*** dans $\mathbb{R}^3$ si $\sigma(\operatorname{Fr}\Omega) < \infty$; cette dernière condition est discutée en détail au chapitre 7. Notons que, dans ce cas,

$$\operatorname{Fr}\Omega = \operatorname{Fr} D_1 \cup \operatorname{Fr} D_2 \cup \cdots \cup \operatorname{Fr} D_n\,;$$

si x est un point régulier de $\operatorname{Fr} D_i$, on définira $\nu(x)$ comme la normale unitaire dirigée vers l'extérieur (par rapport à D_1) si $i = 1$ et comme la normale unitaire

dirigée vers l'intérieur (par rapport à D_i) si $i \geqslant 2$; alors, à chaque point régulier x de $\operatorname{Fr}\Omega$, $\nu(x)$ sera une normale unitaire dirigée vers l'extérieur de Ω dans le sens que, pour un $\varepsilon_0 > 0$,

$$x + \varepsilon\nu(x) \in \operatorname{ext}\Omega\,, \quad x - \varepsilon\nu(x) \in \Omega\,, \quad 0 < \varepsilon < \varepsilon_0\,.$$

Nous dirons que ce choix de la normale unitaire ν dirigée vers l'extérieur de Ω a orienté $\operatorname{Fr}\Omega$ ***positivement*** ; le passage de ν à $-\nu$ partout oriente $\operatorname{Fr}\Omega$ ***négativement***. Si F est un champ vectoriel donné (défini au moins sur $\operatorname{Fr}\Omega$), on note le ***flux*** de F à travers $\operatorname{Fr}\Omega$ (vers l'extérieur de Ω) par

$$\int_{\operatorname{Fr}\Omega} F\cdot\nu\,\mathrm{d}\sigma = \sum_{i=1}^{n}\int_{\operatorname{Fr}D_i} F\cdot\nu\,\mathrm{d}\sigma$$

où $\Omega = D_1\backslash\overline{D}_2\cup\cdots\cup\overline{D}_n$ est un domaine régulier et ν oriente $\operatorname{Fr}\Omega$ positivement.

Pour $F = (F_1, F_2, F_3)$ (défini dans une partie ouverte de $\mathbb{R}^3$), on pose

$$\operatorname{div} F = \mathrm{D}_1F_1 + \mathrm{D}_2F_2 + \mathrm{D}_3F_3\ ;$$

$\operatorname{div} F$ s'appelle la ***divergence*** du champ vectoriel F.

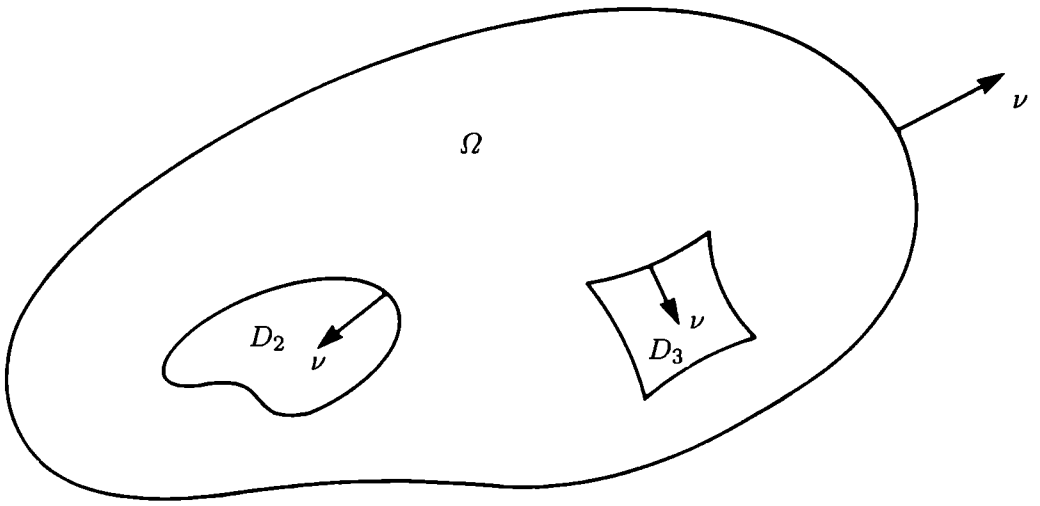


Fig. 6.12 Orientation positive de $\operatorname{Fr}\Omega$.

Nous donnons maintenant une formulation du théorème de Gauss.

6.4.2 Proposition (*théorème de la divergence*)

Soit Ω un domaine régulier dans $\mathbb{R}^3$. Si $F \in C^1(\overline{\Omega};\mathbb{R}^3)$, alors

$$\int_{\operatorname{Fr}\Omega} F\cdot\nu\,\mathrm{d}\sigma = \int_{\Omega}\operatorname{div} F\,\mathrm{d}x \tag{6.24}$$

où $\operatorname{Fr}\Omega$ est orientée positivement par le choix de la normale unitaire ν dirigée vers l'extérieur de Ω.

REMARQUES. Rappelons que, selon nos définitions, ν est définie seulement aux points réguliers de $\operatorname{Fr}\Omega$. Une démonstration d'une version plus générale de cette proposition sera donnée dans le chapitre 7. L'énoncé de la proposition se lit comme suit : *le flux de F à travers $\operatorname{Fr}\Omega$ vers l'extérieur de Ω est égal à l'intégrale volumique de la divergence de F dans Ω.*

On peut donner une interprétation physique de cette proposition en termes du champ de vitesse F d'un écoulement dans le domaine Ω. Elle est exactement la même que celle que nous avons donnée dans le paragraphe 6.2.8 pour un écoulement bidimensionnel en termes du théorème de Green formulé comme un théorème du flux à travers une frontière.

6.4.3 Une vérification du théorème de la divergence

Notons les points de $\mathbb{R}^3$ comme (x, y, z). Supposons que $F = (0, 0, R)$ et que le domaine borné Ω est ***z-simple*** dans le sens que

$$\Omega = \{(x, y, z) : (x, y) \in V\,,\ \alpha(x, y) < z < \beta(x, y)\}\,;$$

ceci veut dire géométriquement que chaque droite verticale d'équation $x = x_0$, $y = y_0$, soit intersecte Ω dans un intervalle ouvert borné (d'extrémités $\alpha(x_0, y_0)$, $\beta(x_0, y_0)$) soit ne le rencontre pas ; l'ensemble V est la projection orthogonale de Ω sur le plan (x, y) de sorte que V est un domaine borné dans $\mathbb{R}^2$. Nous supposons de plus que α, β sont continues dans $\overline{V}$ et même de classe C^1 dans V. Ces conditions ne suffisent pas pour faire de Ω un domaine régulier dans notre sens ; supposons donc, de plus, que $\operatorname{Fr} V$ est une courbe de Jordan régulière par morceaux et que Ω est un domaine régulier. Alors, $\operatorname{Fr}\Omega$ est formée de trois ensembles deux à deux disjoints :

$$S_1 = \{(x, y, z) : (x, y) \in V\,,\ z = \alpha(x, y)\}\,,$$
$$S_2 = \{(x, y, z) : (x, y) \in V\,,\ z = \beta(x, y)\}\,,$$

et un ensemble T formé des points tels que $(x, y) \in \operatorname{Fr} V$. Si (x, y, z) est un point régulier de T avec $z \neq \alpha(x, y)$, $z \neq \beta(x, y)$, alors (x, y, z) est contenu dans un morceau de surface régulière paramétrée par

$$(t, u) \mapsto (\varphi_1(t),\, \varphi_2(t),\, u)\,,$$

$-1 < t < 1$, $z - \varepsilon < u < z + \varepsilon$ où $\varphi_1(0) = x$, $\varphi_2(0) = y$, (φ_1, φ_2) étant un paramétrage d'une partie de $\operatorname{Fr} V$ contenant (x, y) ; un calcul simple montre que ν au point (x, y, z) est de la forme $(\nu_1, \nu_2, 0)$, de sorte que $F \cdot \nu$ au point (x, y, z) est 0. Le flux à travers $\operatorname{Fr}\Omega$ (vers l'extérieur) est donc égal à

$$\iint_V R(x,\, y,\, \beta(x, y))\, dx\, dy - \iint_V R(x,\, y,\, \alpha(x, y))\, dx\, dy$$

selon l'exercice 6 (§ 6.3.12) ; noter que $\nu \cdot (0,0,1) < 0$ sur S_1 et $\nu \cdot (0,0,1) > 0$ sur S_2 ; on en déduit que

$$\int_{\mathrm{Fr}\,\Omega} F \cdot \nu \,\mathrm{d}\sigma = \iint_V \left\{ \int_{\alpha(x,y)}^{\beta(x,y)} \frac{\partial R}{\partial z}(x,y,z)\,\mathrm{d}z \right\} \mathrm{d}x\,\mathrm{d}y$$

$$= \int_\Omega (\mathrm{div}\, F)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\ ;$$

ce qui est l'énoncé du théorème de la divergence dans ce cas avec $F = (0,0,R)$, $\mathrm{div}\, F = \partial R/\partial z$. Si Ω est un domaine régulier qui est x-simple ou y-simple, on obtiendra de même le théorème de la divergence pour $F = (P,0,0)$ ou $F = (0,Q,0)$; finalement, si Ω est un domaine régulier qui est à la fois x, y et z-simple, on obtiendra le théorème de la divergence, par linéarité, pour le cas général de $F = (P,Q,R)$.

Le théorème de la divergence nous permet de calculer facilement le flux à travers une frontière d'un domaine régulier lorsque la divergence du champ est simple à intégrer (par exemple, lorsque la divergence est une constante, éventuellement nulle) ; la situation inverse peut aussi se présenter dans laquelle le flux à travers une frontière est facile à calculer, ce qui donne l'évaluation de l'intégrale de la divergence.

Nous donnons maintenant quelques applications du théorème de la divergence.

6.4.4 Angle solide

Considérons un morceau de surface régulière donné par un paramétrage (ψ, U) ; soit $\overline{V} \subset U$ où V est un domaine tel que $\mathrm{Fr}\, V = C$ est une courbe de Jordan régulière par morceaux ; posons $S = \psi(V)$, $\partial S = \psi(C)$; nous dirons que ∂S est le ***bord*** du morceau de surface régulière S. On vérifie que $\overline{S} = S \cup \partial S$. Supposons que $\overline{S} \subset \mathrm{ext}\, B(0;a)$ et que chaque demi-droite d'origine 0 intersecte $\overline{S}$ à au plus un point ; soit $S(0;a) = \mathrm{Fr}\, B(0;a)$ la sphère de centre 0 et de rayon $a > 0$. Soit $\xi : \overline{S} \to S(0;a)$ l'application bijective définie par

$$\xi(x) = \frac{ax}{\|x\|}\ ;$$

il est clair que ξ est un homéomorphisme entre $\overline{S}$ et $\xi(\overline{S})$ car ξ est continue dans l'ensemble compact $\overline{S}$. Posons $T = \xi(S)$; intuitivement T est « l'ombre » de S sur la sphère $S(0;a)$; on définit l'***angle solide*** formé par S au point 0 comme étant mesuré par la quantité

$$\omega = \frac{\sigma(T)}{a^2}$$

où $\sigma(T)$ est l'aire de la partie T de la sphère $S(0;a)$.

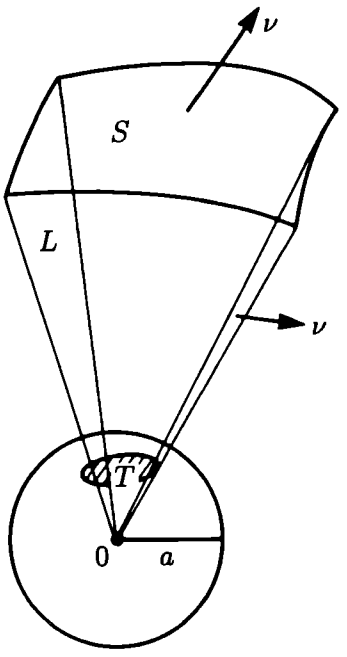


Fig.6.13

En supposant que le domaine Ω formé par la réunion des segments ouverts $]\, x\,,\, \xi(x)\,[$, $x \in S$, est un domaine régulier, nous *démontrons* que

$$\omega = \int_S \frac{x \cdot \nu}{\|x\|^3}\, \mathrm{d}\sigma$$

où ν est la normale unitaire à S dirigée vers l'extérieur de Ω ; en particulier, ω *est indépendant du nombre* a. En effet, si

$$F(x) = \frac{x}{\|x\|^3}\,, \quad x \neq 0\,,$$

on vérifie par un calcul simple que $\operatorname{div} F(x) = 0$ pour tout $x \neq 0$; du théorème de la divergence, on déduit que

$$0 = \int_\Omega \operatorname{div} F\, \mathrm{d}x = \int_S F \cdot \nu\, \mathrm{d}\sigma + \int_T F \cdot \nu\, \mathrm{d}\sigma + \int_L F \cdot \nu\, \mathrm{d}\sigma$$

où L est la partie « latérale » de $\operatorname{Fr}\Omega$ formée par certains segments de type $]\, x\,,\, \xi(x)[$, $x \in \partial S$ et ν est la normale unitaire à $\operatorname{Fr}\Omega$ dirigée vers l'extérieur de Ω. Or, si $y \in\,]\, x\,,\, \xi(x)[$, $x \in \partial S$, ν sera orthogonale à x de sorte que $F \cdot \nu = 0$ sur L car $F(x)$ est proportionnel à x ; d'autre part,

$$\int_T F \cdot \nu\, \mathrm{d}\sigma = \int_T \frac{x}{\|x\|^3} \cdot \frac{-x}{a}\, \mathrm{d}\sigma = -\frac{1}{a^2} \int_T \mathrm{d}\sigma = -\frac{\sigma(T)}{a^2} = -\omega$$

d'où l'on obtient la formule donnée pour ω.

Notons la formule analytique suivante, conséquence immédiate de la formule pour ω :

$$\omega = \left| \iint_U \frac{1}{\|\psi\|^3} [\psi, \mathrm{D}_1\psi, \mathrm{D}_2\psi] \, \mathrm{d}u_1 \, \mathrm{d}u_2 \right|$$

où $[a, b, c]$ est le produit mixte des trois vecteurs a, b, c. De la définition initiale de ω on a $0 \leqslant \omega \leqslant 4\pi$.

6.4.5 Loi de Gauss

Soit

$$F(x) = \frac{kx}{\|x\|^3}, \quad x \in \mathbb{R}^3 \setminus \{0\},$$

où k est un nombre réel non nul. Le champ vectoriel F correspond à la force (d'attraction si $k < 0$, de répulsion si $k > 0$) centrale exercée par une masse ou charge placée à 0 selon la loi d'inverse carré. Nous démontrons que *si Ω est un domaine régulier dans $\mathbb{R}^3$, le flux de F à travers* $\operatorname{Fr}\Omega$ *(vers l'extérieur de Ω) est nul si* $0 \in \operatorname{ext}\Omega$ *et égal à* $4\pi k$ *si* $0 \in \Omega$. En effet, on a déjà vu dans le paragraphe 6.4.4 que $\operatorname{div} F(x) = 0$ pour $x \neq 0$; si $0 \in \operatorname{ext}\Omega$, on peut utiliser le théorème de la divergence directement pour obtenir

$$\int_{\operatorname{Fr}\Omega} F \cdot \nu \, \mathrm{d}\sigma = \int_\Omega \operatorname{div} F \, \mathrm{d}x = 0,$$

ν étant la normale unitaire dirigée vers l'extérieur ; ceci établit notre affirmation pour le cas où $0 \in \operatorname{ext}\Omega$. Si $0 \in \Omega$, prenons un $a > 0$ tel que $\overline{B}(0; a) \subset \Omega$ et soit Ω' le domaine borné régulier $\Omega \setminus \overline{B}(0; a)$; il est clair que

$$\operatorname{Fr}\Omega' = \operatorname{Fr}\Omega \cup S(0; a)$$

où $S(0; a)$ est la sphère de centre 0 et de rayon a. Si ν est la normale unitaire sur $\operatorname{Fr}\Omega'$ dirigée vers l'extérieur de Ω', on a

$$\nu(x) = -\frac{x}{a}, \quad x \in S(0; a).$$

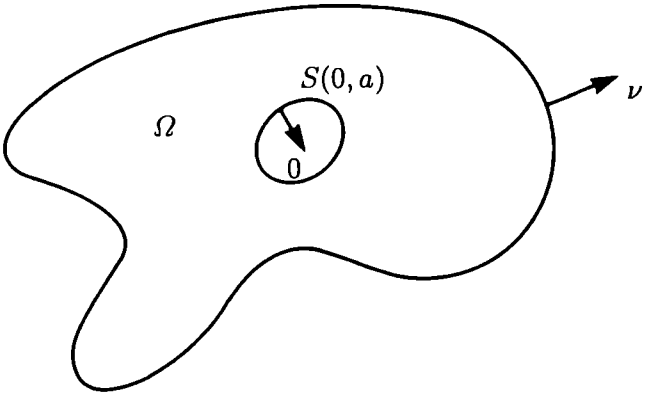


Fig. 6.14

Puisque $0 \in \operatorname{ext} \Omega'$, on a par ce qui précède

$$0 = \int_{\operatorname{Fr} \Omega'} F \cdot \nu \, \mathrm{d}\sigma = \int_{\operatorname{Fr} \Omega} F \cdot \nu \, \mathrm{d}\sigma + \int_{S(0;a)} F \cdot \nu \, \mathrm{d}\sigma \, ;$$

la dernière intégrale vaut

$$\int_{S(0;a)} \frac{kx}{\|x\|^3} \cdot \frac{-x}{a} \, \mathrm{d}\sigma = -\frac{k}{a^2} \int_{S(0;a)} \mathrm{d}\sigma = -4\pi k \, ,$$

ce qui donne

$$\int_{\operatorname{Fr} \Omega} F \cdot \nu \, \mathrm{d}\sigma = 4\pi k$$

et établit notre affirmation.

Ce dernier résultat montre l'importance de la condition suivant laquelle F doit être dans $C^1(\overline{\Omega}; \mathbb{R}^3)$ pour que le théorème de la divergence soit applicable au flux de F à travers $\operatorname{Fr} \Omega$; une seule singularité de F peut rendre son application impossible.

6.4.6 Formule de Green

Soit Ω un domaine régulier dans $\mathbb{R}^3$. *Si f, g sont deux fonctions de classes $C^2(\overline{\Omega})$ alors*

$$\int_{\operatorname{Fr} \Omega} f \mathrm{D}_\nu g \, \mathrm{d}\sigma = \int_\Omega (f \, \Delta g + \nabla f \cdot \nabla g) \, \mathrm{d}x \tag{6.25}$$

où ν est la normale unitaire dirigée vers l'extérieur de Ω.

C'est une conséquence immédiate de (6.24) du théorème de la divergence ; en posant $F = f \, \nabla g$, on obtient

$$F \cdot \nu = f \, \mathrm{D}_\nu g \, , \quad \operatorname{div} F = \operatorname{div}(f \, \nabla g) = f \, \Delta g + \nabla f \cdot \nabla g$$

d'où on déduit (6.25) comme étant un cas particulier de (6.24).

Notons les corollaires suivants de (6.25) :

$$\int_{\operatorname{Fr} \Omega} (f \, \mathrm{D}_\nu g - g \, \mathrm{D}_\nu f) \, \mathrm{d}\sigma = \int_\Omega (f \, \Delta g - g \, \Delta f) \, \mathrm{d}x \tag{6.26}$$

où f, g, ν sont comme pour (6.25). En effet, en écrivant (6.25) avec les rôles de f et g interchangés, on obtient

$$\int_{\operatorname{Fr} \Omega} g \, \mathrm{D}_\nu f \, \mathrm{d}\sigma = \int_\Omega (g \, \Delta f + \nabla f \cdot \nabla g) \, \mathrm{d}x \, ;$$

en soustrayant cette dernière de (6.25), on obtient (6.26). En prenant $f \equiv 1$ dans (6.25), on obtient

$$\int_{\operatorname{Fr} \Omega} \mathrm{D}_\nu g \, \mathrm{d}\sigma = \int_\Omega \Delta g \, \mathrm{d}x \tag{6.27}$$

si g, ν sont comme pour (6.25). En prenant $g = h$, une fonction harmonique dans Ω, de classe $C^2(\overline{\Omega})$, dans (6.27), on obtient

$$\int_{\mathrm{Fr}\,\Omega} \mathrm{D}_\nu h \,\mathrm{d}\sigma = 0\,. \tag{6.28}$$

En prenant dans (6.25) $f \equiv 1$ et $g = h \in C^2(\overline{\Omega})$, h étant une fonction harmonique dans Ω, on obtient

$$\int_{\mathrm{Fr}\,\Omega} h\,\mathrm{D}_\nu h \,\mathrm{d}\sigma = \int_\Omega \|\nabla h\|^2 \,\mathrm{d}x\,. \tag{6.29}$$

Les formules (6.25) et (6.26) s'appellent les ***formules de Green*** dans $\mathbb{R}^3$. Les formules analogues sont valables dans $\mathbb{R}^N$, $N \geqslant 1$; pour $N = 1$ et Ω un intervalle ouvert borné, on obtient ces formules par une intégration par parties, $\mathrm{Fr}\,\Omega$ étant alors réduit à deux points ; nous donnerons la formule exacte pour $N = 1$ dans les chapitres concernant les équations différentielles ; pour $N = 2$, les formules analogues s'obtiennent à partir du théorème de Green dans sa forme du flux à travers une frontière comme donné dans le paragraphe 6.2.8 ; le cas $N > 3$ se déduit du théorème de la divergence dans $\mathbb{R}^N$, $N > 3$, donné dans le chapitre 7 en utilisant les même raisonnements que ceux présentés ici.

On peut obtenir les résultats concernant l'unicité des solutions du problème de Dirichlet en utilisant ces formules ; nous en donnons quelques exemples dans les exercices du paragraphe 6.4.9 ; en général, le principe du maximum (prop. 3.6.1 ou § 6.4.7) donne de meilleurs résultats d'unicité (cf. § 3.6.4).

6.4.7 Propriétés des fonctions harmoniques

Soit f une fonction harmonique dans un domaine D de $\mathbb{R}^3$; rappelons que ceci veut dire que f est de classe C^2 et $\Delta f = 0$ dans D. *Si* $\overline{B}(a;r) \subset D$ alors

$$f(a) = \frac{1}{4\pi r^2} \int_{S(a;r)} f \,\mathrm{d}\sigma \tag{6.30}$$

où $S(a;r) = \{x : \|x - a\| = r\}$ $(r > 0)$.

Pour démontrer (6.30), il suffit de considérer le cas $a = 0$; en écrivant S_r pour $S(0;r)$, S pour $S_r \cup S_\varepsilon$ $(0 < \varepsilon < r)$, $\Omega = \{x : \varepsilon < \|x\| < r\}$, on aura $\mathrm{Fr}\,\Omega = S$; si $g(x) = -1/\|x\|$, $x \neq 0$, on sait que g est harmonique dans $\mathbb{R}^3 \setminus \{0\}$; en utilisant (6.26) on obtient

$$\int_S f\,\mathrm{D}_\nu g \,\mathrm{d}\sigma = \int_S g\,\mathrm{D}_\nu f \,\mathrm{d}\sigma$$

où ν est la normale unitaire dirigée vers l'extérieur de Ω. Or

$$\int_S f\,\mathrm{D}_\nu g \,\mathrm{d}\sigma = \frac{1}{r^2}\int_{S_r} f \,\mathrm{d}\sigma - \frac{1}{\varepsilon^2}\int_{S_\varepsilon} f \,\mathrm{d}\sigma$$

car

$$\mathrm{D}_\nu g(x) = \frac{x}{\|x\|^3} \cdot \frac{x}{r} = \frac{1}{r^2} \quad \text{si } x \in S_r$$

et, de même,

$$\mathrm{D}_\nu g(x) = \frac{x}{\|x\|^3} \cdot \frac{-x}{\varepsilon} = -\frac{1}{\varepsilon^2} \quad \text{si } x \in S_\varepsilon .$$

D'autre part,

$$\begin{aligned}\int_S g\mathrm{D}_\nu f \,\mathrm{d}\sigma &= \int_{S_r} g\mathrm{D}_\nu f \,\mathrm{d}\sigma + \int_{S_\varepsilon} g\mathrm{D}_\nu f \,\mathrm{d}\sigma \\ &= \frac{-1}{r}\int_{S_r} \mathrm{D}_\nu f \,\mathrm{d}\sigma - \frac{1}{\varepsilon}\int_{S_\varepsilon} \mathrm{D}_\nu f \,\mathrm{d}\sigma = 0\end{aligned}$$

selon (6.28) car f est harmonique dans un domaine contenant $\overline{B}(0;r)$. Au total, pour tout ε, $0 < \varepsilon < r$, on a

$$\frac{1}{r^2}\int_{S_r} f \,\mathrm{d}\sigma = \frac{1}{\varepsilon^2}\int_{S_\varepsilon} f \,\mathrm{d}\sigma$$

d'où, en laissant $\varepsilon \to 0$ et en utilisant la continuité de f à 0 et le fait que $\sigma(S_\varepsilon) = 4\pi\varepsilon^2$, on obtient

$$\frac{1}{r^2}\int_{S_r} f \,\mathrm{d}\sigma = 4\pi f(0)$$

ce qui donne (6.30). ♦

La formule (6.30) s'interprète comme suit : *la valeur moyenne d'une fonction harmonique sur une sphère est égale à sa valeur au centre.* On peut transformer (6.30) pour avoir un énoncé concernant la valeur moyenne volumique d'une fonction harmonique ; *avec les notations de (6.30), on a*

$$f(a) = \left(\frac{4}{3}\pi r^3\right)^{-1} \int_{B(a;r)} f(x)\,\mathrm{d}x . \tag{6.31}$$

En effet, de l'exercice 20 (§ 6.3.12), on a

$$\begin{aligned}\int_{B(a;r)} f(x)\,\mathrm{d}x &= \int_0^r \left(\int_{S(a;t)} f \,\mathrm{d}\sigma\right) \mathrm{d}t \\ &= \int_0^r 4\pi t^2 f(a)\,\mathrm{d}t = \frac{4\pi}{3} r^3 f(a)\end{aligned}$$

en utilisant (6.30) ; ceci établit (6.31).

On peut maintenant établir le ***principe du maximum*** pour les fonctions harmoniques (cf. § 3.6.2 et 3.6.3) : *si* $M = \sup_{x \in D} f(x) < \infty$ *et* $M = f(a)$ *pour un point* $a \in D$ *alors* f *est constante dans* D, f *étant harmonique dans le domaine* D *de* $\mathbb{R}^3$.

En effet, l'ensemble

$$E = \{x \in D : f(x) = M\}$$

est non vide et fermé dans D ; E est aussi ouvert dans D car, selon (6.31), si $f(a) = M$, on aura $f(x) = M$ pour x dans $B(a;r)$; pour voir ceci, noter que selon (6.31)

$$\int_{B(a;r)} \{M - f(x)\}\, \mathrm{d}x = 0$$

ce qui donne, avec l'inégalité $M - f(x) \geqslant 0$ et la continuité de f, que $f(x) = M$ pour $x \in B(a;r)$. Puisque D est un domaine, on obtient $D = E$, ce qui veut dire que f est constante dans D.

Puisque $-f$ est harmonique dès que f est harmonique, on obtient immédiatement le ***principe du minimum*** pour les fonctions harmoniques à partir du principe du maximum : *si* f *est harmonique dans un domaine* D *de* $\mathbb{R}^3$, $m = \inf_{x \in D} f(x) > -\infty$ *et* $m = f(a)$ *pour un point* $a \in D$, *alors* f *est constante dans* D.

Les propriétés des fonctions harmoniques données pour les domaines de $\mathbb{R}^3$ restent valables pour les domaines de $\mathbb{R}^N$, $N \geqslant 2$, sans changements essentiels dans leurs démonstrations ; il suffit d'utiliser les formules de Green correspondantes (qui sont exactement celles données dans le paragraphe 6.4.6) et une fonction harmonique appropriée (à la place de g ci-dessus) qui est constante sur les sphères (par exemple, on prend $g(x) = -\ln \|x\|$ dans $\mathbb{R}^2$ et $g(x) = -\|x\|^{-N+2}$ pour $N > 3$).

Finalement, notons que la propriété de valeur moyenne exprimée pour (6.30) caractérise les fonctions harmoniques ; *une fonction continue* f *qui satisfait (6.30) est nécessairement de classe* C^2 *et est harmonique ; en fait, toute fonction harmonique est automatiquement de classe* C^∞.

Nous démontrerons ces propositions dans les chapitres du volume 3 consacrés aux équations aux dérivées partielles.

6.4.8 Poussée d'Archimède

Soit $\overline{\Omega}$ un corps solide immergé dans un liquide statique ; nous supposons que la surface supérieure du liquide est au niveau $x_3 = 0$ et que Ω est un domaine régulier ; si $p(x)$ est la pression du liquide au point x, une loi physique nous dit que la force totale F exercée sur $\overline{\Omega}$ est donnée par

$$F = -\int_{\mathrm{Fr}\,\Omega} p\nu \, \mathrm{d}\sigma$$

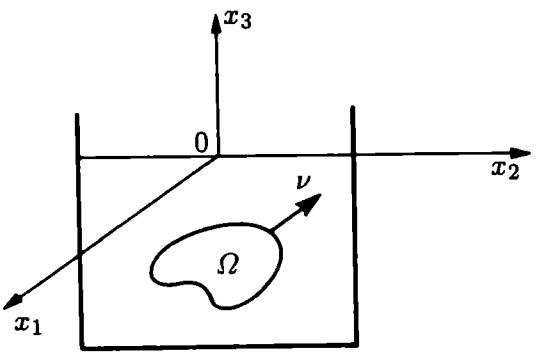


Fig. 6.15

où ν est la normale unitaire dirigée vers l'extérieur de Ω. En supposant que p est de classe C^1 partout, on obtient du théorème de la divergence

$$\int_{\mathrm{Fr}\,\Omega} p\nu \,\mathrm{d}\sigma = \int_{\Omega} (\mathrm{grad}\, p)\,\mathrm{d}x\,.$$

En effet, $p\nu = (p\nu_1, p\nu_2, p\nu_3)$ et

$$\int_{\mathrm{Fr}\,\Omega} p\nu \,\mathrm{d}\sigma = \left(\int_{\mathrm{Fr}\,\Omega} p\nu_1 \,\mathrm{d}\sigma,\ \int_{\mathrm{Fr}\,\Omega} p\nu_2 \,\mathrm{d}\sigma,\ \int_{\mathrm{Fr}\,\Omega} p\nu_3 \,\mathrm{d}\sigma\right);$$

mais

$$\int_{\mathrm{Fr}\,\Omega} p\nu_1 \,\mathrm{d}\sigma = \int_{\mathrm{Fr}\,\Omega} (p,0,0)\cdot\nu \,\mathrm{d}\sigma = \int_{\Omega} (\mathrm{D}_1 p)\,\mathrm{d}x$$

selon le théorème de la divergence et on obtient, de même,

$$\int_{\mathrm{Fr}\,\Omega} p\nu_i \,\mathrm{d}\sigma = \int_{\Omega} (\mathrm{D}_i p)\,\mathrm{d}x\,, \quad i = 1, 2, 3\,,$$

ce qui donne l'égalité voulue pour l'intégrale de $p\nu$ sur $\mathrm{Fr}\,\Omega$. Ainsi, la force totale F est donnée par

$$F = -\int_{\Omega} (\mathrm{grad}\, p)\,\mathrm{d}x\,.$$

Si $p(x) = p_0 - g\rho x_3$, $x = (x_1, x_2, x_3)$, p_0 étant la pression atmosphérique (constante), ρ la densité constante du liquide, g la constante gravitationnelle pour la terre, on en déduit que

$$F = \left(\int_{\Omega} g\rho \,\mathrm{d}x\right) e_3 = (gM)\, e_3$$

où $e_3 = (0,0,1)$ et M est la masse de $\overline{\Omega}$ (égale à la masse de Ω). L'égalité $F = gM\, e_3$ peut être interprétée comme suit : *la force totale sur le corps s'exerce verticalement (vers le haut) et sa magnitude est égale au poids d'un volume du liquide égal au volume du corps.*

6.4.9 Exercices

1 Calculer le flux du champ F à travers S dans les cas suivants ; les points de $\mathbb{R}^3$ sont notés (x, y, z) :

(i) $F(x,y,z) = (2x^2yz,\ -xy^2z,\ -xyz^2)$; S est la sphère centrée en $(0,0,0)$ de rayon $a > 0$ orientée positivement ;

(ii) $F(x,y,z) = (x^3, y^3, z^3)$; S comme dans (i) ;

(iii) $F(x,y,z) = (x,y,z)$; S est la frontière du cube $[0\,,\,1]^3$, orientée positivement ;

(iv) $F(x,y,z) = (xy,\ yz,\ xz)$; S est la pyramide formée par les plans d'équation $x = 0$, $y = 0$, $z = 0$, $x + y + z = 1$, orientée positivement ;

(v) $F(x,y,z) = (y^2, x^2, z)$; S est l'ellipsoïde d'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

orienté positivement (a, b, $c > 0$).

2 Soit Ω un domaine régulier dans $\mathbb{R}^3$; si $F \in C^1(\overline{\Omega}; \mathbb{R}^3)$ est tel que $\operatorname{div} F$ est une constante k, démontrer que

$$\int_{\operatorname{Fr}\Omega} F \cdot \nu \, \mathrm{d}\sigma = k\lambda_3(\Omega)$$

où ν est la normale unitaire dirigée vers l'extérieur de Ω. En particulier, la dernière formule s'applique si

$$F(x) = (x_1, 0, 0) \quad \text{ou si} \quad F(x) = (0, x_2, 0) \quad \text{ou si} \quad F(x) = (0, 0, x_3)$$

ou si $F(x) = x$.

3 Si Ω, ν sont comme dans l'exercice 2 et $F(x_1, x_2, x_3) = (x_1^2,\ x_1x_2,\ x_1x_3)$, démontrer que

$$\frac{1}{4\lambda_3(\Omega)} \int_{\operatorname{Fr}\Omega} F \cdot \nu \, \mathrm{d}\theta = \overline{x}_1$$

où $(\overline{x}_1, \overline{x}_2, \overline{x}_3)$ est le barycentre de Ω.

4 Si Ω, ν sont comme dans l'exercice 2 et $F \in C^2(\overline{\Omega}; \mathbb{R}^3)$, démontrer que

$$\int_{\operatorname{Fr}\Omega} \operatorname{rot} F \cdot \nu \, \mathrm{d}\sigma = 0$$

(Vérifier d'abord que $\operatorname{div}\operatorname{rot} F = 0$; voir (6.34) qui donne le résultat sous une condition plus faible sur F.)

5 Soient

$$\begin{aligned} F(x,y,z) &= (z^2,\, yx^2,\, y^2z) \\ G(x,y,z) &= (x^2y,\, z,\, -x^2z) \end{aligned} \qquad (x,y,z) \in \mathbb{R}^3 .$$

Si Ω est un domaine régulier dans $\mathbb{R}^3$, démontrer que le flux de F à travers $\operatorname{Fr}\Omega$ (vers l'extérieur de Ω) est strictement plus grand que le flux de G à travers $\operatorname{Fr}\Omega$.

6 Pour $F \in C^1(\Omega;\mathbb{R}^3)$, Ω ouvert dans $\mathbb{R}^3$, posons $\operatorname{Flux}(F\,;\,a,r)$ pour le flux de F à travers $S(a;r)$ vers l'extérieur de $B(a;r)$ où $\overline{B}(a;r) \subset \Omega$. Si F, G sont deux champs vectoriels de classe C^1 dans un domaine Ω de $\mathbb{R}^3$, démontrer que $\operatorname{Flux}(F\,;\,a,r) \geqslant \operatorname{Flux}(G\,;\,a,r)$ pour toutes les possibilités de a et r si et seulement si $\operatorname{div} F(x) \geqslant \operatorname{div} G(x)$, $x \in \Omega$.

7 Soit $F \in C^1(A;\mathbb{R}^3)$, A étant un ensemble ouvert dans $\mathbb{R}^3$; si Ω est un domaine régulier tel que $\overline{\Omega} \subset A$, posons $\operatorname{Flux}(F;\Omega)$ pour le flux de F à travers $\operatorname{Fr}\Omega$ vers l'extérieur de Ω. Démontrer que, si $a \in A$,

$$\operatorname{div} F(a) = \lim_{n\to\infty} \frac{1}{\lambda_3(\Omega_n)} \operatorname{Flux}(F\,;\,\operatorname{Fr}\Omega_n)$$

où $\{\Omega_n\}$ est une suite de domaines bornés réguliers tels que $a \in \overline{\Omega}_n$, $\overline{\Omega}_n \subset D$, $n = 1, 2, \ldots$, et diamètre $(\Omega_n) \to 0$ lorsque $n \to \infty$.

8 Soit

$$u(x) = \sum_{i=1}^{n} \frac{q_i}{4\pi\|x - \xi_i\|}, \quad x \in \mathbb{R}^3 \setminus \{\xi_1, \ldots, \xi_n\}$$

où $\xi_1, \ldots, \xi_n$ sont n points dans $\mathbb{R}^3$ et $q_1, \ldots, q_n$ sont des nombres réels (« charges ») ; soit $F = -\nabla u$. Si Ω est un domaine régulier dans $\mathbb{R}^3$ tel que $\operatorname{Fr}\Omega$ ne contient aucun des points $\xi_1, \ldots, \xi_n$, démontrer que le flux de F à travers $\operatorname{Fr}\Omega$ (vers l'extérieur de Ω) est égal à $q_{i_1} + \cdots + q_{i_k}$ si

$$\Omega \cap \{\xi_1, \ldots, \xi_n\} = \{\xi_{i_1}, \ldots, \xi_{i_k}\}$$

et égal à 0 si Ω est disjoint de $\{\xi_1, \ldots, \xi_n\}$ (cf. § 6.4.5).

9 Soit Ω un domaine régulier dans $\mathbb{R}^3$, ν la normale unitaire dirigée vers l'extérieur définie sur $\operatorname{Fr}\Omega$. Démontrer que l'équation

$$\triangle u = f \text{ dans } \Omega\,, \quad \mathrm{D}_\nu u = 0 \text{ sur } \operatorname{Fr}\Omega$$

ne peut avoir une solution u dans $C^2(\overline{\Omega})$ que si

$$\int_\Omega f(x)\,\mathrm{d}x = 0$$

(cf. (6.27), § 6.4.6) ; si u existe, démontrer que u est unique.

10 Soient Ω, ν comme dans l'exercice 9. Si f et g sont deux fonctions dans $C^2(\overline{\Omega})$ telles que $\Delta f = af$, $\Delta g = ag$ dans Ω où $a : \Omega \to \mathbb{R}$ est une fonction quelconque, alors démontrer que

$$\int_{\mathrm{Fr}\,\Omega} (f\,\mathrm{D}_\nu g - g\,\mathrm{D}_\nu f)\,\mathrm{d}\sigma = 0\,.$$

11 Soit $f \in C^2(A)$, A étant un ensemble ouvert dans $\mathbb{R}^3$. Démontrer que, pour $a \in A$,

$$\Delta f(a) = \lim_{n\to\infty} \frac{1}{\lambda_3(\Omega_n)} \int_{\mathrm{Fr}\,\Omega_n} \mathrm{D}_\nu f\,\mathrm{d}\sigma$$

où $\{\Omega_n\}$ est une suite de domaines réguliers tels que $a \in \Omega_n$, $\overline{\Omega}_n \subset A$, $n = 1, 2, \ldots$, et diamètre$(\Omega_n) \to 0$ lorsque $n \to \infty$, ν étant la normale unitaire dirigée vers l'extérieur (cf. (6.27), § 6.4.6).

6.5 THÉORÈME DE STOKES DANS $\mathbb{R}^3$

6.5.1 Introduction

Dans cette section, nous présentons une seconde généralisation à $\mathbb{R}^3$ du théorème de Green dans $\mathbb{R}^2$ (prop. 6.2.5) ; rappelons que le théorème de la divergence (prop. 6.4.2) est aussi une généralisation à $\mathbb{R}^3$ du théorème de Green. Un cas particulier important du théorème de Stokes dans $\mathbb{R}^3$ concerne l'évaluation d'une intégrale curviligne (appelée la ***circulation*** du champ vectoriel F le long de γ)

$$\int_\gamma F \cdot \mathrm{d}\vec{s}$$

où γ est un lacet régulier par morceaux situé dans une surface régulière orientée S dans $\mathbb{R}^3$; si γ forme le « bord » d'un « domaine » $\Omega \subset S$ et si γ est orienté d'une manière appropriée, alors on a selon le théorème de Stokes

$$\int_\gamma F \cdot \mathrm{d}\vec{s} = \int_\Omega \mathrm{rot}\,F \cdot \nu\,\mathrm{d}\sigma \tag{6.32}$$

où ν est la normale unitaire qui donne l'orientation de S.

La formule (6.32) est lue de manière suivante : *la circulation d'un champ vectoriel F le long d'un lacet γ est égale au flux du rotationnel de F à travers Ω* où Ω est la partie de la surface S dont le « bord » est donné par γ.

L'orientation de γ est décrite intuitivement comme suit : si l'on parcourt le « bord » de Ω suivant l'orientation de γ, alors Ω est toujours à notre « gauche » si nous nous déplaçons avec notre tête dans la direction indiquée par ν.

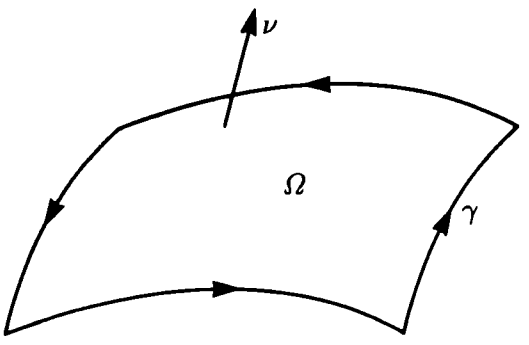


Fig. 6.16

Nous formulons maintenant ces considérations intuitives de manière précise.

6.5.2 Domaines réguliers et leurs bords dans une surface

Soit S une surface régulière dans $\mathbb{R}^3$. Nous fixons d'abord quelques notations pour les considérations topologiques relatives à S (cf. sect. 1.5). Soit $A \subset S$; on dit que A est ouvert dans S ou S-ouvert si $A = V \cap S$, V étant ouvert dans $\mathbb{R}^3$; un voisinage S-ouvert de $x \in S$ est un ensemble S-ouvert contenant x. Notons par $\mathrm{Fr}_S A$ la frontière de A relative à S ; $x \in \mathrm{Fr}_S A$ si et seulement si $x \in S$ et pour tout voisinage S-ouvert B de x, on a

$$B \cap A \neq \emptyset, \quad B \cap (S \setminus A) \neq \emptyset .$$

On note par $\overline{A}_S$ l'adhérence de A dans S ; on a alors

$$\overline{A}_S = A \cup \mathrm{Fr}_S A .$$

Soit C une courbe de Jordan régulière par morceaux dans la surface S et D une partie de S, ouverte dans S, telle que $\mathrm{Fr}_S D = C$; on dira que D est un ***domaine régulier*** (***simple***) dans S et que C est le ***bord*** de D, en symboles,

$$\partial D = C .$$

Une partie Ω de S s'appellera un ***domaine régulier*** dans S si

$$\Omega = D_1 \setminus (\overline{D}_2 \cup \cdots \cup \overline{D}_n)$$

où chaque D_i est un domaine régulier simple dans S, $\overline{D}_i \subset D_1$ si $i = 2, \ldots, n$, et $\overline{D}_2, \ldots, \overline{D}_n$ sont deux à deux disjoints. On posera

$$\partial\Omega = \bigcup_{i=1}^{n} \partial D_i ;$$

$\partial\Omega$ s'appelle le ***bord*** de Ω ; il est clair que $\partial\Omega = \mathrm{Fr}_S \Omega$; donc le bord de Ω est la réunion de n courbes de Jordan régulières par morceaux $C_1, \ldots, C_n$, deux à deux disjointes, avec $C_i = \partial D_i$, $i = 1, \ldots, n$, les courbes $C_2, \ldots, C_n$ étant dans le « domaine intérieur » D_1 défini par C_1.

Si S est une surface régulière compacte dans $\mathbb{R}^3$ (elle est alors automatiquement orientable), on pose $\partial S = \emptyset$; on verra dans le chapitre 7 le sens géométrique de cette convention. Ainsi, si S est une sphère dans $\mathbb{R}^3$, son bord ∂S est $\emptyset$. Une surface régulière compacte S dans $\mathbb{R}^3$ est la frontière d'un domaine borné V dans $\mathbb{R}^3$ (selon un théorème de Brouwer déjà mentionné) et on choisit l'orientation (dite ***positive***) de S en prenant la normale unitaire ν dirigée vers l'extérieur de V. Ces considérations se généralisent aux surfaces compactes S régulières par morceaux ; elles sont aussi frontières de domaines bornés et on posera $\partial S = \emptyset$ pour ces surfaces compactes S régulières par morceaux ; pour une telle surface S, l'ensemble S tout entier sera considéré comme un domaine régulier de S ayant un bord vide.

6.5.3 Orientation du bord

Soit S une surface régulière dans $\mathbb{R}^3$ orientée par le choix d'une normale unitaire continue ν. Dans ce paragraphe, nous définissons la notion de l'orientation ***positive*** (par rapport à l'orientation de S fixée par ν) du bord d'un domaine régulier dans S.

Prenons d'abord un domaine régulier simple D dans S avec $\partial D = C$; soit x un point régulier de C. Soit U un disque centré en $(0,0) \in \mathbb{R}^2$ et $\psi : U \to \mathbb{R}^3$ un paramétrage d'un morceau de surface régulière M contenant x, $M \subset S$, tel que $\psi(0,0) = x$ et tel que ν sur M est la normalisation de $\mathrm{D}_1\psi \times \mathrm{D}_2\psi$; posons

$$U_+ = \psi^{-1}(M \cap D), \quad U_- = \psi^{-1}\big(M \cap (\overline{D})^c\big), \quad \Gamma = \psi^{-1}(M \cap C).$$

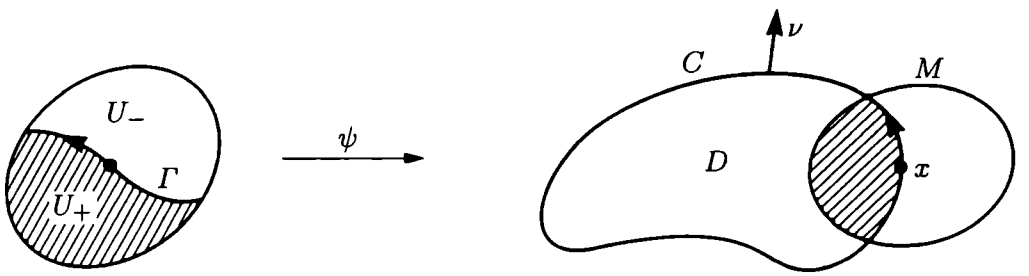


Fig. 6.17

L'orientation ***positive*** de la partie de C dans M est celle induite par ψ de Γ orientée positivement par rapport à U_+ ; ceci donne l'orientation locale positive autour de chaque point régulier de C. On peut démontrer que ces orientations locales positives donnent une orientation positive de C globalement. L'orientation ***négative*** de C est celle donnée par un parcours opposé. Ces considérations seront reprises plus précisément dans le chapitre 7.

Une façon analytique de décrire ce qui précède est la suivante. Soit γ un paramétrage approprié de C, une courbe de Jordan régulière par morceaux et x un point régulier de C, $x = \gamma(t)$, $\gamma'(t) \neq 0$; le vecteur $\tau = \gamma'(t)$ est dans T_x, l'espace tangent à la surface S au point x. Soit $\xi \in T_x$, ξ, τ étant linéairement

indépendants (rappelons que ξ, τ sont des éléments de $\mathbb{R}^3$) ; supposons que ξ est tel qu'il existe un chemin régulier

$$\beta : [0, \varepsilon_0] \longrightarrow S$$

avec $\beta(0) = x$, $\beta(]0, \varepsilon_0[) \subset D$, $\beta'(0) = \xi$; nous dirons que ξ est un vecteur de T_x dirigé vers l'***intérieur*** de D. Nous verrons dans la suite que de tels vecteurs ξ existent à chaque point régulier x de C ; si ν est un multiple positif de $\tau \times \xi$ pour tout point régulier x de C, nous pourrons conclure que C est orientée ***positivement*** par γ. On verra la relation de ce critère avec la première définition dans le paragraphe qui suit.

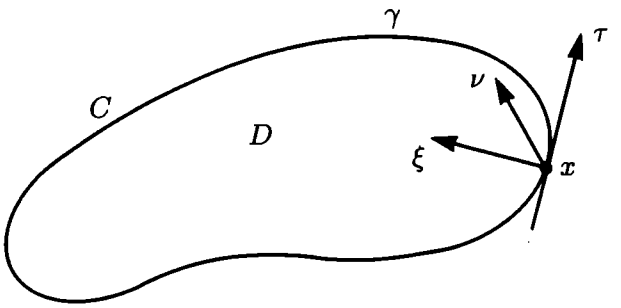


Fig. 6.18

Si $\tau \times \xi$ est un multiple négatif de ν pour tout point régulier x de C, nous dirons que C est orientée ***négativement*** par γ ; il est clair que si C est orientée positivement par γ, alors le paramétrage opposé $\widetilde{\gamma}$ oriente C négativement.

Si Ω est un domaine régulier dans S donné comme

$$\Omega = D_1 \setminus (\overline{D}_2 \cup \cdots \cup \overline{D}_n)$$

(avec $D_1, \ldots, D_n$ comme spécifiés dans le paragraphe 6.5.2), alors un paramètrage $\gamma = \gamma_1 \vee \cdots \vee \gamma_n$ de $\partial\Omega$ (qui consiste de n courbes $C_1, \ldots, C_n$ régulières par morceaux, deux à deux disjointes) donne une orientation ***positive*** à $\partial\Omega$ (par rapport à l'orientation de S fixée par ν) si γ_1 oriente $C_1 = \partial D_1$ positivement et γ_i oriente $C_i = \partial D_i$, $i = 2, \ldots, n$, négativement ; si γ_1 oriente ∂D_1 négativement et γ_i oriente ∂D_i, $i = 2, \ldots, n$, positivement, on dira que γ oriente $\partial\Omega$ ***négativement***.

Après ces présentations, nous pouvons énoncer le théorème de Stokes précisément dans $\mathbb{R}^3$.

6.5.4 Proposition (*théorème de Stokes dans* $\mathbb{R}^3$)

Soit S une surface régulière dans $\mathbb{R}^3$ orientée par le choix d'une normale unitaire continue ν. Soit Ω un domaine régulier dans S ; si F est un champ vectoriel dans $C^1(\overline{\Omega}; \mathbb{R}^3)$ alors

$$\int_{\partial\Omega} F \cdot \mathrm{d}\vec{s} = \int_{\Omega} \operatorname{rot} F \cdot \nu \, \mathrm{d}\sigma \tag{6.33}$$

où $\partial\Omega$ est orienté positivement.

REMARQUES. Explicitons un cas particulier de (6.33) contenu dans notre énoncé ; si S est une surface *compacte*, régulière par morceaux, dans $\mathbb{R}^3$, S est considérée comme un domaine régulier de lui-même avec $\partial S = \emptyset$; si F est un champ vectoriel de classe C^1, défini dans une partie ouverte contenant S, alors

$$\int_S \operatorname{rot} F \cdot \nu \, \mathrm{d}\sigma = 0 . \tag{6.34}$$

Notons que (6.33) contient le théorème de Green dans $\mathbb{R}^2$; pour le voir, on prend S comme

$$\{(x_1, x_2, 0) : (x_1, x_2) \in \mathbb{R}^2\} \subset \mathbb{R}^3$$

orientée par $\nu(x_1, x_2, 0) = (0, 0, 1)$; si Ω est un domaine régulier dans S, alors Ω considéré comme une partie de $\mathbb{R}^2$ est aussi un domaine régulier de $\mathbb{R}^2$ et on peut considérer $\partial\Omega$ comme la frontière de Ω dans cette interprétation ; si F est un champ vectoriel de la forme

$$F(x_1, x_2, x_3) = \big(P(x_1, x_2), Q(x_1, x_2), 0\big)$$

alors

$$\operatorname{rot} F(x_1, x_2, x_3) = \big(0, 0, \mathrm{D}_1 Q(x_1, x_2) - \mathrm{D}_2 P(x_1, x_2)\big)$$

$$\operatorname{rot} F \cdot \nu = \mathrm{D}_1 Q - \mathrm{D}_2 P$$

$$\int_{\partial\Omega} F \cdot \mathrm{d}\vec{s} = \int_{\operatorname{Fr}\Omega} (P \, \mathrm{d}x_1 + Q \, \mathrm{d}x_2)$$

$$\int_\Omega \operatorname{rot} F \cdot \nu \, \mathrm{d}\sigma = \int_\Omega (\mathrm{D}_1 Q - \mathrm{D}_2 P) \, \mathrm{d}x_1 \, \mathrm{d}x_2$$

d'où l'on remarque que (6.33) est, dans ce cas, rien d'autre que le théorème de Green.

Une démonstration détaillée d'un théorème plus général que le théorème de Stokes dans $\mathbb{R}^3$ sera donnée dans le chapitre 7. Pour l'immédiat, nous nous contentons d'une vérification de cette proposition dans un cas particulier simple où, en fait, on la déduit du théorème de Green dans $\mathbb{R}^2$.

6.5.5 Une vérification du théorème de Stokes

La vérification concerne un morceau de surface régulière S défini par un paramétrage $\psi : U \to \mathbb{R}^3$, orienté par le choix de la normale unitaire continue donné par

$$\nu = (\mathrm{D}_1\psi \times \mathrm{D}_2\psi) / \|\mathrm{D}_1\psi \times \mathrm{D}_2\psi\|$$

et un domaine régulier simple $\Omega \subset S$ donné comme l'image par ψ d'un domaine régulier V de $\mathbb{R}^2$ tel que $\overline{V} \subset U$; nous posons C pour la courbe de Jordan régulière par morceaux qui forme $\operatorname{Fr} V$ de sorte que

$$\Omega = \psi(V) , \quad \partial\Omega = \psi(C) .$$

Nous supposons que C est paramétré positivement (dans $\mathbb{R}^2$) par le lacet simple régulier par morceaux $\alpha : I \to \mathbb{R}^2$ et que $\partial\Omega$ est paramétré par $\gamma = \psi \circ \alpha$. Vérifions d'abord que γ donne une orientation positive à $\partial\Omega$, la surface S étant orientée par ν. En effet, soit x un point régulier de $\partial\Omega$ et soit $u \in C$ tel que $\psi(u) = x$; soit $\alpha(t) = u$, $t \in I$; il est clair que $\alpha'(t) \neq 0$ car $\gamma'(x) = \psi'(u)\alpha'(t) \neq 0$.

Soit $a = \alpha'(t)$; soit b le vecteur normal unitaire au point u de C tel que b est dirigée vers l'intérieur de V de sorte que $\psi(u + \varepsilon b)$ est dans Ω pour $\varepsilon > 0$ suffisamment petit ; la courbe définie par

$$\varepsilon \mapsto \psi(u + \varepsilon b) = \beta(\varepsilon)$$

est telle que $\beta'(0) = \psi'(u)b$. Posons $M = \psi'(u)$; on a donc $\gamma'(x) = Ma$ et Mb qui sont deux vecteurs linéairement indépendants de T_x, espace tangent à S au point x, avec Mb un vecteur dirigé vers l'intérieur de Ω ; si $a = (a_1, a_2)$, $b = (b_1, b_2)$, on a

$$Ma = a_1 \, \mathrm{D}_1\psi + a_2 \, \mathrm{D}_2\psi \,, \quad Mb = b_1 \, \mathrm{D}_1\psi + b_2 \, \mathrm{D}_2\psi$$

($\mathrm{D}_1\psi$, $\mathrm{D}_2\psi$ calculés au point u) d'où

$$Ma \times Mb = (a_1 b_2 - a_2 b_1)(\mathrm{D}_1\psi \times \mathrm{D}_2\psi) \,;$$

puisque $(a_1 b_2 - a_2 b_1) > 0$ (cf. § 6.2.4), on en déduit que γ oriente $\partial\Omega$ positivement par rapport à l'orientation fixée par ν (cf. § 6.5.3).

Nous allons vérifier (6.33) pour un champ de type $F = (F_1, 0, 0)$ en faisant l'hypothèse supplémentaire que le paramétrage ψ est de classe C^2 ; une démonstration analogue est possible pour $F = (0, F_2, 0)$ ou $F = (0, 0, F_3)$, ce qui donne (6.33) pour un F général par linéarité. Dans le cas $F = (F_1, 0, 0)$, (6.33) devient

$$\int_{\partial\Omega} F_1 \, \mathrm{d}x_1 = \int_{\Omega} \{\mathrm{D}_3 F_1 \, \mathrm{d}x_3 \wedge \mathrm{d}x_1 - \mathrm{D}_2 F_1 \, \mathrm{d}x_1 \wedge \mathrm{d}x_2\} \,. \tag{6.35}$$

Le principe de démonstration de (6.35) est de transformer le côté gauche, en nous servant de ψ, en une intégrale curviligne dans U où nous nous servons du théorème de Green pour obtenir ensuite l'expression à droite de (6.35). Ecrivons

$$u = (u_1, u_2) \,, \quad \psi = (\psi_1, \psi_2, \psi_3) \,,$$
$$\alpha = (\alpha_1, \alpha_2) : [a \,, b] \longrightarrow \mathbb{R}^2 \,, \quad \gamma = (\gamma_1, \gamma_2, \gamma_3)$$

où $\gamma = \psi \circ \alpha$, $\gamma_1 = \psi_1 \circ \alpha$. Alors

$$\int_{\partial\Omega} F_1 . \mathrm{d}x_1 = \int_{\alpha} \{g_1 \, \mathrm{d}u_1 + g_2 \, \mathrm{d}u_2\}$$

avec

$$g_1 = F_1 \circ \psi \cdot \mathrm{D}_1\psi_1 \,, \quad g_2 = F_1 \circ \psi \cdot \mathrm{D}_2\psi_1$$

car

$$F_1\,\mathrm{d}x_1 = F_1 \circ \gamma \cdot \gamma_1'\,\mathrm{d}t = (F_1 \circ \psi \circ \alpha)(\mathrm{D}_1\psi_1 \circ \alpha \cdot \alpha_1' + \mathrm{D}_2\psi_1 \circ \alpha \cdot \alpha_2')\,\mathrm{d}t$$

d'où, selon le théorème de Green dans $\mathbb{R}^2$,

$$\int_{\partial\Omega} F_1\,\mathrm{d}x_1 = \iint_V \{\mathrm{D}_1 g_2 - \mathrm{D}_2 g_1\}\,\mathrm{d}u_1\,\mathrm{d}u_2\,. \tag{6.36}$$

A partir de

$$\begin{aligned}\mathrm{D}_1 g_2 = {} & \{\mathrm{D}_1 F_1 \circ \psi \cdot \mathrm{D}_1\psi_1 + \mathrm{D}_2 F_1 \circ \psi \cdot \mathrm{D}_1\psi_2 + \mathrm{D}_3 F_1 \circ \psi \cdot \mathrm{D}_1\psi_3\}\mathrm{D}_2\psi_1 \\ & + F_1 \circ \psi \cdot D_{12}\psi_1\end{aligned}$$

$$\begin{aligned}\mathrm{D}_2 g_1 = {} & \{\mathrm{D}_1 F_1 \circ \psi \cdot \mathrm{D}_2\psi_1 + \mathrm{D}_2 F_1 \circ \psi \cdot \mathrm{D}_2\psi_2 + \mathrm{D}_3 F_1 \circ \psi \cdot \mathrm{D}_2\psi_3\}\mathrm{D}_1\psi_1 \\ & + F_1 \circ \psi \cdot D_{12}\psi_1\end{aligned}$$

on obtient

$$\mathrm{D}_1 g_2 - \mathrm{D}_2 g_1 = (\mathrm{D}_3 F_1 \circ \psi)\frac{\partial(\psi_3, \psi_1)}{\partial(u_1, u_2)} - (\mathrm{D}_2 F_1 \circ \psi)\frac{\partial(\psi_1, \psi_2)}{\partial(u_1, u_2)}\,.$$

En introduisant ceci dans le côté droit de (6.36), on obtient (6.35) compte tenu des notations introduites dans le paragraphe 6.3.6.

Le cas où le paramétrage ψ est seulement de classe C^1 se déduit du cas précédent par approximation ; nous omettons les détails de cette approximation (donnée, par exemple, dans de Rham [34, p. 33]).

6.5.6 Remarques

Une série d'applications de (6.33) (théorème de Stokes dans $\mathbb{R}^3$) provient du fait qu'une des deux intégrales qui y figurent peut être facile à calculer dans une situation donnée. La formule (6.33) donne aussi un sens physique au vecteur $\operatorname{rot} F$; supposons que F soit interprété comme un champ de vitesse d'un fluide en mouvement ; considérons Ω comme un disque de rayon r, centré à un point $a \in \mathbb{R}^3$ et contenu dans un plan de $\mathbb{R}^3$ ayant la normale unitaire ν (un vecteur constant) ; alors $\partial\Omega$ est un cercle C_r que l'on peut imaginer comme le bord d'une petite roue légère Ω centrée à a et ayant une tendance à tourner autour d'un axe passant par a dans la direction fixée par ν. En supposant que F est de classe C^1, on déduit de (6.33) (en orientant C_r correctement) que

$$\operatorname{rot} F(a) \cdot \nu = \lim_{r \to 0} \frac{1}{\pi r^2} \int_{C_r} F \cdot \mathrm{d}\vec{s}\,;$$

on interprète cette égalité en disant que la composante de $\operatorname{rot} F(a)$ dans la direction ν est égale à la circulation par unité d'aire, au point a, dans un plan ayant la normale unitaire ν. Si $\operatorname{rot} F(a) = 0$ pour tout point a dans un domaine D, on dit que F est ***irrotationnel*** dans D ; on interprète ceci en disant que le fluide n'a pas un effet de rotation dans D. Si $\operatorname{rot} F(a) \neq 0$ pour tout a dans un domaine D, on constatera un effet rotatif dans l'écoulement du fluide dans D.

Notons une autre particularité de la formule (6.33). Une courbe de Jordan régulière par morceaux C peut être considérée comme le bord $\partial\Omega$ de différents domaines Ω contenus dans différentes surfaces ; ainsi, un cercle C peut être considéré comme le bord d'un domaine planaire (un disque) ou le bord d'une calotte d'une sphère quelconque contenant C ; l'intégrale curviligne d'un champ F sur C (orienté) ne dépend pas du choix du domaine Ω dont C est le bord. Ce fait donne quelquefois des égalités intéressantes provenant de l'intégrale de flux à travers Ω dans (6.33).

6.5.7 Exercices

1 Calculer $\int_C F \cdot \mathrm{d}\vec{s}$, en utilisant le théorème de Stokes, dans les exemples suivants :

(i) $F(x,y,z) = (z, -x, -y)$; C est le chemin triangulaire de sommets $(a,0,0)$, $(0,a,0)$, $(0,0,a)$ parcouru dans cet ordre (avec $a > 0$) ;

(ii) $F(x,y,z) = (z-y,\ x-z,\ y-x)$; C est un cercle de rayon r dans le plan d'équation $x+y+z=1$ orienté de telle manière que l'intégrale est $\geqslant 0$;

(iii) $F(x,y,z) = a(x)(1,1,1)$, $a \in C^1(\mathbb{R})$; C est le cercle dans le plan d'équation $x-y-z=1$ de centre $(0,0,-1)$ et rayon 1, orienté positivement par rapport au choix de la normale $(1,-1,-1)$ du plan.

(iv) $F(x,y,z) = \big(f(x),\ g(x,y,z),\ -g(x,y,z)\big)$, f et g étant de classe C^1 partout ; C est une courbe de Jordan régulière par morceaux dans le plan d'équation $y-z=0$ orienté par la normale $(0,1,-1)$.

(v) $F(x,y,z) = (y,z,x)$; C est la courbe obtenue par l'intersection des surfaces d'équations $x+y=1$ et $x^2+y^2+z^2 = x+y$ respectivement, orientée de sorte que l'intégrale est positive.

2 Soit $F(x) = \omega \times x$, $x \in \mathbb{R}^3$, ω étant constant ; vérifier que $\operatorname{rot} F(x) = 2\omega$. En conclure que

$$\int_C F \cdot \mathrm{d}\vec{s} = 2\omega \cdot \nu\, \sigma(D)$$

si C est une courbe de Jordan régulière par morceaux dans un plan P de $\mathbb{R}^3$ ayant une normale unitaire ν, D est le domaine intérieur dans le plan P déterminé par C et C est orientée convenablement.

3 Soit $F(x) = Ax$, $x \in \mathbb{R}^3$, $A = [a_{ij}]$ une matrice constante du type $(3,3)$; vérifier que

$$\operatorname{rot} F(x) = (a_{32} - a_{23},\ a_{13} - a_{31},\ a_{21} - a_{12}) \cdot$$

Obtenir une formule pour $\int_C F \cdot \mathrm{d}\vec{s}$ où C est comme dans l'exercice 2.

4 Soit

$$F(x,y,z) = \left(\frac{-y}{x^2+y^2}\,,\ \frac{x}{x^2+y^2}\,,\ f(x) \right)\,, \quad (x,y,z) \in \Omega\,,$$

où $\Omega = \{(x,y,z) : x^2 + y^2 > 0\}$ et $f \in C^1(\mathbb{R})$. Vérifier que :

(i) $\operatorname{rot} F = 0$ dans Ω,
(ii) $\int_C F \cdot \mathrm{d}\vec{s} = 2\pi$,

si C est un cercle dans le plan d'équation $z = 0$ de centre $(0,0,0)$ et de rayon r orienté positivement. Donc F est irrotationnel mais non conservatif.

5 Soit

$$G(x,y) = \left(\frac{-y}{x^2+ky^2}\,,\ \frac{x}{x^2+ky^2}\,,\ 0 \right), \quad (x,y) \neq (0,0)\,,\ k > 0\,;$$

démontrer qu'il n'existe pas de champ vectoriel F de classe C^1 dans $\Omega = \mathbb{R}^3 \setminus \{(0,0,0)\}$ tel que $\operatorname{rot} F = 0$ dans Ω et $F(x,y,0) = G(x,y)$ si $(x,y) \neq (0,0)$. (Si C est l'ellipse d'équation $x^2 + ky^2 = 1$ dans le plan $z = 0$, considérer $I = \int_C (-y\,\mathrm{d}x + x\,\mathrm{d}y)$, C étant orientée de manière appropriée ; l'existence de F donnerait $I = 0$ en contradiction avec $I > 0$.)

6 Démontrer les formules suivantes à partir des théorème de Stokes et de Gauss, en les interprétant correctement :

(i) $\displaystyle\int_{\operatorname{Fr}\Omega} f\nu\,\mathrm{d}\sigma = \int_\Omega (\operatorname{grad} f)\,\mathrm{d}x\,;$

(ii) $\displaystyle\int_{\operatorname{Fr}\Omega} (F \times \nu)\,\mathrm{d}\sigma = -\int_\Omega (\operatorname{rot} F)\,\mathrm{d}x\,;$

(iii) $\displaystyle\int_{\partial D} f\tau\,\mathrm{d}s = -\int_D (\operatorname{grad} f) \times \nu\,\mathrm{d}\sigma.$

Dans (i), (ii), Ω est un domaine régulier dans $\mathbb{R}^3$ et ν est la normale unitaire sur $\operatorname{Fr}\Omega$ dirigée vers l'extérieur ; (i) et (ii) sont des conséquences du théorème de Gauss ; par exemple, en calculant la première composante dans (ii), on a

$$\begin{aligned}\int_{\operatorname{Fr}\Omega} (F_2\nu_3 - F_3\nu_2)\,\mathrm{d}\sigma &= \int_{\operatorname{Fr}\Omega} G \cdot \nu\,\mathrm{d}\sigma \quad \text{où } G = (0, -F_3, F_2) \\ &= \int_\Omega (\operatorname{div} G)\,\mathrm{d}x = -\int_\Omega (\mathrm{D}_2F_3 - \mathrm{D}_3F_2)\,\mathrm{d}x\,.\end{aligned}$$

Dans (iii), D est un domaine régulier dans une surface orientée par la normale unitaire continue ν et τ est le vecteur tangent unitaire de ∂D orienté comme il faut ; en calculant la première composante du côté gauche, on aura

$$\begin{aligned}\int_{\partial D} f\,\mathrm{d}x_1 &= \int_{\partial D} F \cdot \mathrm{d}\vec{s} \quad \text{où } F = (f,0,0) \\ &= \int_D (\operatorname{rot} F \cdot \nu)\,\mathrm{d}\sigma = -\int_D \{(\mathrm{D}_2 f)\nu_3 - (\mathrm{D}_3 f)\nu_2\}\,\mathrm{d}\sigma\,.\end{aligned}$$

[7] Soit F un champ vectoriel de classe C^1 dans un domaine V de $\mathbb{R}^3$. En utilisant l'exercice 6(ii), démontrer que, pour tout $a \in V$,

$$\operatorname{rot} F(a) = \lim_{n\to\infty} \frac{1}{\lambda_3(\Omega_n)} \int_{\operatorname{Fr}\Omega_n} (\nu \times F)\, \mathrm{d}\sigma$$

où $\{\Omega_n\}$ est une suite de domaines réguliers dans V avec $a \in \overline{\Omega}_n$, diamètre $(\Omega_n) \to 0$ $(n \to \infty)$, ν étant chaque fois la normale unitaire dirigée vers l'extérieur.

[8] En suivant la discussion du paragraphe 6.5.6, formuler et démontrer la proposition suivante :

$$\nu \cdot \operatorname{rot} F(a) = \lim_{n\to\infty} \frac{1}{\beta_n} \int_{C_n} F \cdot \mathrm{d}\vec{s}$$

si $\{C_n\}$ est une suite de courbes appropriées, convenablement orientées, ν est un vecteur unitaire donné et $\{\beta_n\}$ une suite de nombres appropriés (« aire enfermée par C_n »).

6.6 QUELQUES PROBLÈMES CONCERNANT LES CHAMPS VECTORIELS DANS $\mathbb{R}^3$

Tous les champs vectoriels considérés dans cette section sont définis dans une partie ouverte Ω de $\mathbb{R}^3$; il s'agit donc des fonctions $F : \Omega \to \mathbb{R}^3$. Nous commençons par une énumération des formules qui sont souvent utiles en analyse vectorielle.

6.6.1 Formules concernant les champs vectoriels

Dans les formules suivantes, α et β sont dans $\mathbb{R}$, f et g sont des fonctions scalaires de type $\Omega \to \mathbb{R}$, F et G sont des champs vectoriels dans Ω ; f, g, F, G sont supposés être C^1 dans les formules (1) à (9) ; dans les autres formules, nous supposons qu'ils sont C^2 :

(1) $\operatorname{grad}(\alpha f + \beta g) = \alpha \operatorname{grad} f + \beta \operatorname{grad} g$;
(2) $\operatorname{grad}(fg) = f \operatorname{grad} g + g \operatorname{grad} f$;
(3) $\operatorname{div}(\alpha F + \beta G) = \alpha \operatorname{div} F + \beta \operatorname{div} G$;
(4) $\operatorname{rot}(\alpha F + \beta G) = \alpha \operatorname{rot} F + \beta \operatorname{rot} G$;
(5) $\operatorname{grad}(F \cdot G) = (F \cdot \nabla)G + (G \cdot \nabla)F + F \times \operatorname{rot} G + G \times \operatorname{rot} F$;
(6) $\operatorname{div}(fF) = f \operatorname{div} F + (\operatorname{grad} f) \cdot F$;
(7) $\operatorname{div}(F \times G) = (\operatorname{rot} F) \cdot G - F \cdot (\operatorname{rot} G)$;
(8) $\operatorname{rot}(fF) = f \operatorname{rot} F + (\operatorname{grad} f) \times F$;
(9) $\operatorname{rot}(F \times G) = (\operatorname{div} G)F - (\operatorname{div} F)G + (G \cdot \nabla)F - (F \cdot \nabla)G$;

(10) $\operatorname{div}\operatorname{grad} f = \Delta f$;
(11) $\operatorname{rot}\operatorname{grad} f = 0$;
(12) $\operatorname{div}\operatorname{rot} F = 0$;
(13) $\operatorname{rot}\operatorname{rot} F = \operatorname{grad}\operatorname{div} F - \Delta F$;
(14) $\operatorname{div}(\nabla f \times \nabla g) = 0$;
(15) $\operatorname{div}(f\,\nabla g - g\,\nabla f) = f\,\Delta g - g\,\Delta f$.

On a utilisé ∇ *et* grad *interchangeablement ; le sens de* $(F \cdot \nabla)G$ *est le suivant : si* $G = (G_1, , G_2, G_3)$ *alors*

$$(F \cdot \nabla)G = (F \cdot \nabla G_1,\ F \cdot \nabla G_2,\ F \cdot \nabla G_3)\ ;$$

si $F = (F_1, F_2, F_3)$, $\Delta F = (\Delta F_1, \Delta F_2, \Delta F_3)$ *où* Δ *est l'opérateur laplacien* $(\Delta f = \mathrm{D}_{11} f + \mathrm{D}_{22} f + \mathrm{D}_{33} f)$.

La vérification de ces formules se fait par des calculs directs. Par exemple, pour établir (13), comparons les composantes des deux côtés ; si $F = (F_1, F_2, F_3)$

$$\begin{aligned}\big(\operatorname{rot}\operatorname{rot} F\big)_1 &= \mathrm{D}_2(\mathrm{D}_1 F_2 - \mathrm{D}_2 F_1) - \mathrm{D}_3(\mathrm{D}_3 F_1 - \mathrm{D}_1 F_3)\\ &= \mathrm{D}_{12} F_2 + \mathrm{D}_{13} F_3 - (\mathrm{D}_{22} F_1 + \mathrm{D}_{33} F_1)\end{aligned}$$

$$\begin{aligned}\big(\operatorname{grad}\operatorname{div} F\big)_1 - \big(\Delta F\big)_1 &= \mathrm{D}_1(\mathrm{D}_1 F_1 + \mathrm{D}_2 F_2 + \mathrm{D}_3 F_3) - \Delta F_1\\ &= \mathrm{D}_{12} F_2 + \mathrm{D}_{13} F_3 - (\Delta F_1 - \mathrm{D}_{11} F_1)\end{aligned}$$

d'où l'on obtient l'égalité entre les premières composantes des deux côtés de (13).

Nous laissons les vérifications des formules (1) à (15) comme exercices. Les formules (11), (12) et (13) jouent un rôle important dans les paragraphes suivants de cette section.

Ajoutons quelques remarques d'ordre mnémonique concernant les notations. Le symbole ∇ (lu *nabla*, le nom d'un ancien instrument de musique, en forme de triangle renversé, cf. Chilov [9, p. 325]) manipulé formellement comme un vecteur ordinaire de $\mathbb{R}^3$ permet de récupérer un certain nombre de formules (1) à (15) (mais pas toutes) ; ainsi en écrivant $\nabla = (\mathrm{D}_1, \mathrm{D}_2, \mathrm{D}_3)$, $F = (F_1, F_2, F_3)$, on obtient bien

$$\nabla f = \operatorname{grad} f\,, \quad \nabla \cdot F = \operatorname{div} F\,, \quad \nabla \times F = \operatorname{rot} F\ ;$$

de plus,

$$\nabla \times \nabla f = 0\,, \quad \nabla \cdot \nabla \times F = 0$$

correspondent aux formules (11) et (12) et sont faciles à retenir en analogie avec les formules vectorielles élémentaires

$$a \times a = 0\,, \quad a \cdot a \times b = 0\ ;$$

de même, la formule vectorielle

$$a \times (b \times c) = b(a \cdot c) - c(a \cdot b) = (a \cdot c)b - (a \cdot b)c$$

donne par analogie

$$\operatorname{rot}\operatorname{rot} F = \nabla \times (\nabla \times F) = \nabla(\nabla \cdot F) - (\nabla \cdot \nabla)F$$

ce qui est la formule (13), compte tenu d'une autre analogie,

$$\nabla \cdot \nabla = \Delta$$

donnant la formule (10) ; plusieurs auteurs écrivent ∇^2 pour Δ. Si nous écrivons

$$\nabla = \mathrm{D}_1 e_1 + \mathrm{D}_2 e_2 + \mathrm{D}_3 e_3$$
$$F = F_1 e_1 + F_2 e_2 + F_3 e_3$$

et si nous utilisons les règles de multiplication des *quaternions* de Hamilton données par

$$e_1 e_2 = -e_2 e_1 = e_3\,, \quad e_2 e_3 = -e_3 e_2 = e_1\,, \quad e_3 e_1 = -e_1 e_3 = e_2$$
$$e_1^2 = e_2^2 = e_3^2 = -1$$

nous obtenons formellement :

$$\begin{aligned}\nabla F &= (\mathrm{D}_1 e_1 + \mathrm{D}_2 e_2 + \mathrm{D}_3 e_3)(F_1 e_1 + F_2 e_2 + F_3 e_3)\\ &= -(\mathrm{D}_1 F_1 + \mathrm{D}_2 F_2 + \mathrm{D}_3 F_3)\\ &\quad + (\mathrm{D}_2 F_3 - \mathrm{D}_3 F_2) e_1 + (\mathrm{D}_3 F_1 - \mathrm{D}_1 F_3) e_2 + (\mathrm{D}_1 F_2 - \mathrm{D}_2 F_1) e_3\\ &= -\operatorname{div} F + \operatorname{rot} F\,.\end{aligned}$$

Ceci a conduit plusieurs auteurs à appeler ∇ l'*opérateur hamiltonien* ; nous n'utiliserons pas cette terminologie car l'opérateur hamiltonien a un sens bien différent dans plusieurs autres branches de physique mathématique (surtout en mécanique quantique).

6.6.2 Champ irrotationnel, champ incompressible

Soit $F \in C^1(\Omega; \mathbb{R}^3)$; on dit que F est ***irrotationnel*** si $\operatorname{rot} F = 0$ dans Ω ; F s'appelle ***incompressible*** ou ***solénoïdal*** si $\operatorname{div} F = 0$ dans Ω. De la formule (11) (§ 6.6.1), on voit qu'un champ F de la forme $\operatorname{grad} u$, $u \in C^2(\Omega)$, dérivant d'un ***potentiel scalaire*** u, est irrotationnel. De la formule (12) (§ 6.6.1), on déduit qu'un champ F de la forme $\operatorname{rot} A$, $A \in C^2(\Omega; \mathbb{R}^3)$, dérivant d'un ***potentiel vecteur*** A, est incompressible. La proposition suivante dit, entre autre chose, que les réciproques de ces deux observations sont valables si Ω est un domaine étoilé ; on verra dans la discussion qui suit qu'une condition géométrique sur Ω est bien nécessaire pour pouvoir affirmer ces réciproques.

6.6.3 Proposition

Soient Ω un domaine de $\mathbb{R}^3$ et F un champ vectoriel de classe C^1 dans Ω.

(i) Si $F = \operatorname{grad} u$, $u \in C^2(\Omega)$ alors F est irrotationnel dans Ω ; si F est irrotationnel dans Ω et Ω est un domaine étoilé, alors il existe $u \in C^2(\Omega)$ tel que $F = \operatorname{grad} u$ dans Ω ; le potentiel scalaire u est unique à une constante additive près.

(ii) Si $F = \operatorname{rot} A$, $A \in C^2(\Omega; \mathbb{R}^3)$, alors F est incompressible dans Ω ; si F est incompressible dans Ω et Ω est un domaine étoilé, alors il existe $A \in C^2(\Omega; \mathbb{R}^3)$ tel que $F = \operatorname{rot} A$ dans Ω ; le potentiel vecteur A est unique à l'addition près d'un champ dérivant d'un potentiel scalaire.

(iii) Pour un domaine Ω étoilé, F est à la fois irrotationnel et incompressible si et seulement si $F = \operatorname{grad} u$ où u est une fonction harmonique dans Ω ; u est unique à une constante additive près.

DÉMONSTRATION

(i) On a déjà vu que si $F = \operatorname{grad} u$, $u \in C^2(\Omega)$, alors $\operatorname{rot} F = 0$ dans Ω, même si Ω est un ensemble ouvert quelconque dans $\mathbb{R}^3$; l'unicité de u (à une constante additive près) vient de la connexité de Ω observée dans le paragraphe 6.1.9 ; l'existence de u tel que $F = \operatorname{grad} u$ lorsque $\operatorname{rot} F = 0$ dans Ω, un domaine étoilé, était démontré dans la proposition 6.1.15 dans le cadre plus général de $\mathbb{R}^N$. Rappelons que si Ω est étoilé par rapport à $0 \in \Omega$ alors une formule pour u est la suivante :

$$u(x) = \int_0^1 F(tx) \cdot x \, \mathrm{d}t = \int_{[0,x]} F \cdot \mathrm{d}\vec{s}, \quad x \in \Omega .$$

(ii) Le fait que $F = \operatorname{rot} A$, $A \in C^2(\Omega; \mathbb{R}^3)$, entraîne que $\operatorname{div} F = 0$ dans Ω a déjà été vu et ici Ω peut être une partie ouverte quelconque ; démontrons d'abord le degré d'unicité indiqué dans le choix de A si Ω est un domaine étoilé ; si $B = A + \operatorname{grad} u$, $u \in C^2(\Omega)$, alors il est évident de (i) que si $F = \operatorname{rot} A$ dans Ω, on aura aussi

$$\operatorname{rot} B = \operatorname{rot} A + \operatorname{rot} \operatorname{grad} u = F$$

dans Ω, Ω étant une partie ouverte quelconque. D'autre part, si, pour A, B dans $C^2(\Omega; \mathbb{R}^3)$,

$$F = \operatorname{rot} A = \operatorname{rot} B$$

dans Ω un domaine étoilé, alors

$$\operatorname{rot}(B - A) = 0$$

dans Ω d'où, selon (i), $B = A + \operatorname{grad} u$, u étant de classe C^2.

Soit maintenant $\operatorname{div} F = 0$ dans le domaine étoilé Ω ; sans perte de généralité, nous pouvons supposer que Ω est étoilé par rapport 0 que nous supposons être dans Ω (voir le raisonnement dans la démonstration de la proposition 6.1.15). Posons

$$A(x) = \int_0^1 \{F(tx) \times tx\}\, dt, \quad x \in \Omega .$$

Vérifions que $\operatorname{rot} A = F$ dans Ω. Si $A = (A_1, A_2, A_3)$, $F = (F_1, F_2, F_3)$, $x = (x_1, x_2, x_3)$, on aura pour la première composante de $\operatorname{rot} A$,

$$(\operatorname{rot} A)_1 = D_2 A_3 - D_3 A_2 = \int_0^1 \Phi(t, x)\, dt$$

où

$$\Phi(t, x) = D_2 \left[t\{F_1(tx)x_2 - F_2(tx)x_1\}\right] - D_3 \left[t\{F_3(tx)x_1 - F_1(tx)x_3\}\right] .$$

L'échange de la dérivation et de l'intégration est facilement justifiable à cause de la continuité des dérivées partielles de F_1, F_2, F_3. Un calcul montre que

$$\begin{aligned}\Phi(t, x) &= 2tF_1(tx) + t^2\{x_2 D_2 F_1(tx) + x_3 D_3 F_1(tx)\} - t^2 x_1\{D_2 F_2(tx) + D_3 F_3(tx)\} \\ &= 2tF_1(tx) + t^2\{x_1 D_1 F_1(tx) + x_2 D_2 F_1(tx) + x_3 D_3 F_1(tx)\} \\ &= 2tF_1(tx) + t^2 \frac{d}{dt} F_1(tx) = \frac{d}{dt}\{t^2 F_1(tx)\}\end{aligned}$$

en utilisant la condition $\operatorname{div} F = 0$. Ainsi

$$(\operatorname{rot} A)_1 = \int_0^1 \frac{d}{dt}\{t^2 F_1(tx)\}\, dt = F_1(x)$$

et de même pour les deux autres composantes de $\operatorname{rot} A$.

(iii) Si $F = \operatorname{grad} u$, u harmonique dans Ω, alors $\operatorname{rot} F = \operatorname{rot}\operatorname{grad} u = 0$ et $\operatorname{div} F = \Delta u = 0$ dans Ω (Ω ici étant une partie ouverte quelconque). Si Ω est un domaine étoilé et F est tel que

$$\operatorname{rot} F = 0, \quad \operatorname{div} F = 0$$

dans Ω, alors on a tout d'abord $F = \operatorname{grad} u$, $u \in C^2(\Omega)$, grâce à (i) et ensuite

$$0 = \operatorname{div} F = \operatorname{div}\operatorname{grad} u = \Delta u$$

dans Ω, ce qui montre que le potentiel scalaire u est harmonique dans Ω. ♦

6.6.4 Remarques

Soit $\Omega_1 = \mathbb{R}^3 \setminus L$, $L = \{(0,0,t) : t \in \mathbb{R}\}$; on a déjà vu dans l'exercice 4 (§ 6.5.7) que le champ vectoriel

$$F(x_1,x_2,x_3) = \left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2}, 0\right), \quad x = (x_1, x_2, x_3) \in \Omega_1,$$

est irrotationnel mais non conservatif dans Ω_1, de sorte que F ne peut pas être de la forme $\operatorname{grad} u$. Noter que F est de classe C^∞ et que Ω_1 est non étoilé. On peut interpréter F physiquement comme le champ de vitesse d'un fluide qui a une source de tourbillon le long de L (la ligne de singularité).

Soit $\Omega_2 = \mathbb{R}^3 \setminus \{0\}$ et $F(x) = x/\|x\|^3$, $x \in \Omega_2$; on a déjà vu par des calculs simples que, pour $x \in \Omega_2$,

$$F(x) = \operatorname{grad} u(x), \quad u(x) = \frac{-1}{\|x\|}, \quad \Delta u(x) = 0$$

de sorte que F est un champ vectoriel de classe C^∞ qui est à la fois irrotationnel et incompressible dans Ω_2. Néanmoins, F n'est pas de la forme $\operatorname{rot} A$ dans Ω_2 ; en effet, on a déjà vu dans le paragraphe 6.4.5 que

$$\int_S F \cdot \nu \, d\sigma = 4\pi \neq 0$$

si S est une sphère quelconque de centre 0 orientée par sa normale unitaire ν dirigée vers l'extérieur ; d'autre part, on sait que si $F = \operatorname{rot} A$, on aura selon (6.34) du théorème de Stokes

$$\int_S F \cdot \nu \, d\sigma = 0$$

ce qui contredit le calcul précédent.

Ces contre-exemples montrent qu'une condition sur le domaine Ω est bien nécessaire pour la validité de (i) et (ii) de la proposition 6.6.3. Noter que dans les deux contre-exemples, il y a des potentiels locaux ; puisque une boule ouverte quelconque est convexe donc étoilée, il est clair que pour une boule ouverte $B \subset \Omega_1$ (donc B n'est pas « percée » par L), il existe toujours un potentiel scalaire u_B d'où dérive le champ vectoriel restreint à B ; seulement, ces potentiels u_B ne peuvent pas être rassemblés pour construire une seule fonction u définie dans Ω_1 (qui est « univoque » comme l'on disait quelquefois). C'est le problème d'existence globale du potentiel scalaire contre lequel il y a une « obstruction » topologique dans Ω_1. La situation concernant le contre-exemple dans Ω_2 est analogue ; dans chaque boule ouverte $B \subset \Omega_2$ (c'est-à-dire telle que 0 n'est pas dans B), on a un potentiel vecteur A_B tel que $\operatorname{rot} A_B = F|_B$ mais ces potentiels vecteurs A_B ne forment pas un seul potentiel vecteur dans Ω_2. Il est aussi clair que « l'obstruction » dans Ω_1 est d'une nature différente de

celle dans Ω_2. En fait, dans Ω_2, tout champ irrotationnel dérive d'un potentiel scalaire et dans Ω_1, tout champ incompressible dérive d'un potentiel vecteur ; ces affirmations sont immédiates si l'on connaît les caractérisations topologiques des domaines Ω pour lesquels les énoncés (i) ou (ii) de la proposition 6.6.3 sont valables. On reviendra à cette question dans le chapitre 7.

Les domaines étoilés sont les domaines simples mais néanmoins assez généraux pour lesquels on peut établir la proposition 6.6.3. Les domaines étoilés dans $\mathbb{R}^3$ sont homéomorphes à $\mathbb{R}^3$ (ce qui est géométriquement évident mais dont la démonstration n'est pas immédiate ; cf. Berger [2, tome 2, p. 27]) et, en fait, la proposition 6.6.3 est valable pour tout domaine homéomorphe à $\mathbb{R}^3$; mais ce dernier théorème (conséquence d'un théorème célèbre de de Rham) sort du cadre de notre discussion ; un énoncé d'une version simple du théorème de de Rham est donné dans le chapitre 7.

6.6.5 Proposition (*théorème de Helmholtz*)

Soit $F \in C^1(\Omega;\mathbb{R}^3)$ tel que, si $f = \operatorname{div} F$, l'équation $\Delta u = f$ dans Ω possède une solution $u \in C^2(\Omega)$. Alors $F = G + H$ où G est un champ irrotationnel et H est un champ incompressible.

DÉMONSTRATION. Posons $G = \operatorname{grad} u$ où $u \in C^2(\Omega)$ est tel que $\Delta u = f = \operatorname{div} F$ dans Ω ; l'existence de u est garantie par hypothèse. Si $H = F - \operatorname{grad} u$ on voit que

$$\operatorname{div} H = \operatorname{div} F - \operatorname{div}\operatorname{grad} u = f - \Delta u = 0$$

et $F = G + H$, $G = \operatorname{grad} u$, est une décomposition de F demandée car $\operatorname{rot} G = \operatorname{rot}\operatorname{grad} u = 0$. ♦

6.6.6 Remarques

L'équation $\Delta u = f$ dans Ω s'appelle ***équation de Poisson*** (écrite souvent comme $\Delta u = -4\pi\rho$) dans Ω. *Si $f \in C^1(\Omega)$ et bornée et Ω est un ensemble ouvert borné alors*

$$u(x) = \frac{-1}{4\pi} \int_\Omega \frac{f(y)}{\|y - x\|}\, dy \tag{6.37}$$

définit $u \in C^1(\mathbb{R}^3)$ telle que u est de classe C^∞ dans $\operatorname{ext}\Omega$ et de classe C^2 dans Ω avec

$$\Delta u = f \text{ dans } \Omega\,, \quad \Delta u = 0 \text{ dans } \operatorname{ext}\Omega\,.$$

Nous ne démontrons pas cette proposition ici ; en fait, on peut démontrer l'existence d'une solution de l'équation de Poisson dans Ω sous des conditions beaucoup plus faibles sur f et Ω mais la continuité seule de f ne suffit pas (à moins que l'on veuille *généraliser* la notion de solution). Néanmoins, il est connu que l'équation de Poisson $\Delta u = f$ dans Ω possède une solution dès que $f \in C^1(\Omega)$ (et une solution *généralisée* dès que f est continue) ; cette solution n'est jamais unique car si u est une solution et v est une fonction harmonique

quelconque $u + v$ sera aussi une solution ($\Delta(u+v) = \Delta u = f$ dans Ω). Ainsi, la décomposition affirmée par le théorème de Helmholtz ne devient unique que si l'on impose d'autres restrictions sur les champs G et H. Ces questions seront étudiées ultérieurement dans le chapitre du volume 3 sur les équations aux dérivées partielles.

La fonction u définie par (6.37) s'appelle le ***potentiel newtonien*** défini par f. Une étude détaillée des potentiels newtoniens est donnée dans Dautray-Lions [12, vol. 2]. On démontre en fait que, dès que u est bien définie par f, on aura $\Delta u = f$ dans Ω, au moins dans le sens généralisé (de la théorie des distributions).

6.6.7 Exercices

[1] Soit $F(x) = \varphi(\|x\|)x$, $x \in \mathbb{R}^3 \setminus \{0\}$, où $\varphi \in C^1(]0, \infty[)$; vérifier que, si $x \neq 0$, $\operatorname{div} F(x) = 3\varphi(\|x\|) + \|x\|\,\varphi'(\|x\|)$ et $\operatorname{rot} F(x) = 0$; $\operatorname{div} F(x) = 0$ pour tout $x \neq 0$ si et seulement si $\varphi(t) = c/t^3$ (c est une constante réelle). (Noter que F dérive d'un potentiel dès que φ est continue (Ex. 14, § 6.1.18).)

[2] Soit F comme dans l'exercice 1 avec $\varphi(R) \neq 0$. Voir que le flux de F à travers la sphère S_R de centre 0 et rayon R vers l'extérieur est égal à $R\varphi(R)\sigma(S_R) \neq 0$; en déduire que F est un exemple d'un champ qui n'est pas de la forme $\operatorname{rot} A$.

[3] Vérifier les formules suivantes (F étant un champ vectoriel dérivable, $a \in \mathbb{R}^3$ constant) :

(i) $a \cdot \operatorname{rot} F = -\operatorname{div}(a \times F)$;
(ii) $a \times \operatorname{rot} F = \operatorname{grad}(a \cdot F) - (a \cdot \nabla)F$;
(iii) $\operatorname{rot}(a \times F) = (\operatorname{div} F)a - (a \cdot \nabla)F$;
(iv) $\|a\|^2 \operatorname{div} F = a \cdot \{\operatorname{grad}(a \cdot F) + \operatorname{rot}(a \times F)\}$.

[4] Vérifier que

$$\int_{\partial D} \operatorname{rot} F \cdot d\vec{s} = -\int_D (\Delta F \cdot \nu)\, d\sigma \quad \text{si } \operatorname{div} F \text{ est constante}$$

où F est un champ vectoriel de classe C^2 dans un domaine Ω de $\mathbb{R}^3$ et D est un domaine régulier dans une surface régulière dans Ω orientée par la normale unitaire ν.

[5] Soit Ω un domaine régulier dans $\mathbb{R}^3$ tel que $\operatorname{Fr}\Omega$ est donné par une équation $u = 0$, u étant une fonction de classe C^1. Démontrer que

$$\int_{\operatorname{Fr}\Omega} (\operatorname{rot} F) \cdot \frac{\nabla u}{\|\nabla u\|}\, d\sigma = 0$$

où $F \in C^1(\overline{\Omega}; \mathbb{R}^3)$.

6 Soit

$$F(x) = \left(\frac{-x_2}{x_1^2 + x_2^2}, \frac{x_1}{x_1^2 + x_2^2}, 0 \right), \quad x = (x_1, x_2, x_3) \in \Omega,$$

où $\Omega = \mathbb{R}^3 \setminus \{(0,0,t) : t \in \mathbb{R}\}$. Vérifier que $u(x) = \operatorname{arc\,tg}(x_2/x_1)$, $x \in \Omega_1 = \{(x_1, x_2, x_3) : x_1 \neq 0\}$ ainsi que $v(x) = \operatorname{arc\,ctg}(x_1/x_2)$, $x \in \Omega_2 = \{(x_1, x_2, x_3) : x_2 \neq 0\}$ sont des potentiels scalaires pour $F|_{\Omega_1}$ et $F|_{\Omega_2}$ respectivement, $\Omega = \Omega_1 \cup \Omega_2$, mais F ne possède pas de potentiel scalaire dans Ω. Sur $\Omega_1 \cap \Omega_2$, si $x_1 x_2 > 0$ on a $v(x) = u(x)$ et si $x_1 x_2 < 0$ on a $v(x) = u(x) + \pi$.

7 Soit Ω un pavé ouvert dans $\mathbb{R}^3$ contenant 0 ; si $F \in C^1(\Omega; \mathbb{R}^3)$ est tel que $\operatorname{div} F = 0$ dans Ω, vérifier que $F = \operatorname{rot} A$ dans Ω où $A = (A_1, A_2, 0)$ avec

$$A_2(x) = -\int_0^{x_3} F_1(x_1, x_2, t)\, dt$$

$$A_1(x) = \int_0^{x_3} F_2(x_1, x_2, t)\, dt - \int_0^{x_2} F_3(x_1, t, 0)\, dt$$

où $x = (x_1, x_2, x_3)$, $F = (F_1, F_2, F_3)$. (Ceci donne une autre construction du potentiel vecteur dans le cas où le domaine est un pavé.)

8 Soit Ω un domaine dans $\mathbb{R}^3$ étoilé par rapport au point $p \in \Omega$. Si $F \in C^1(\Omega; \mathbb{R}^3)$ avec $\operatorname{div} F = 0$ dans Ω, vérifier à partir de la démonstration de la proposition 6.6.3 que

$$A(x) = \int_0^1 \{F(t(x-p)+p) \times t(x-p)\}\, dt, \quad x \in \Omega,$$

vérifie $\operatorname{rot} A(x) = F(x)$, $x \in \Omega$.

9 Soit

$$F(x) = \frac{x}{\|x\|^3}, \quad x \in \Omega = \mathbb{R}^3 \setminus \{0\}.$$

On sait que $\operatorname{div} F = 0$ dans Ω et que F n'est pas de la forme $\operatorname{rot} A$ dans Ω. Les calculs suivants donnent A tel que $F = \operatorname{rot} A$ dans certains sous-domaines étoilés de Ω. Soit $p \in \mathbb{R}^3$, $\|p\| = 1$; posons

$$\Omega_p = \mathbb{R}^3 \setminus \{tp : t \leqslant 0\} = \{x \in \mathbb{R}^3 : p \cdot x + \|x\| \neq 0\}.$$

Vérifier que Ω_p est étoilé par rapport à p ; la formule pour $A(x)$ donnée dans l'exercice 8 conduit à l'expression

$$A_p(x) = \varphi_p(x)(p \times x), \quad \varphi_p(x) = \int_0^1 \frac{t\, dt}{\|t(x-p)+p\|^3}.$$

On obtient $\varphi_p(x) = 1/\{\|x\|(\|x\| + p.x)\}$ en faisant les calculs suivants :

$$\|t(x-p)+p\|^2 = at^2 + 2bt + c, \quad a = \|x-p\|^2,\ b = p\cdot x - 1,\ c = 1\ ;$$

une primitive de $t/(at^2+2bt+c)^{3/2}$ est

$$\frac{-1}{ac-b^2}\cdot\frac{bt+c}{\sqrt{at^2+2bt+c}}\cdot$$

Noter que si p et q sont linéairement indépendants, $\|p\| = 1$, $\|q\| = 1$, alors $\Omega_p \cup \Omega_q = \Omega$.

10 Vérifier que A_p de l'exercice 9 peut être écrit comme suit, pour $x \in \Omega_p$,

$$A_p(x) = p \times \operatorname{grad} \ln(\|x\| + p\cdot x) = -\operatorname{rot}\{\ln(\|x\| + p\cdot x)p\}\ .$$

11 Etudier si les champs suivants sont de la forme rot A (où $x = (x_1, x_2, x_3) \in \mathbb{R}^3$) :

(i) $F(x) = (x_1^2,\ x_2^2,\ x_3^2)$;
(ii) $F(x) = (x_1^2 + x_2,\ x_3^2 - 2x_1x_2,\ x_2^2)$.

Dans le cas affirmatif, trouver un $A = (A_1, A_2, A_3)$ approprié, par la méthode de l'exercice 7 et par celle de l'exercice 8 et comparer vos résultats. Comment peut-on assurer l'existence d'un tel A avec $A_3(x) \neq 0$ pour tout x ?

12 Soit $F(x) = \big(f(x_2),\ g(x_1),\ h(x)\big)$, $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, où f, g, h sont de classe C^1. Déterminer toutes les fonctions h telles que F soit de la forme rot A (pour f, g données).

13 Soit $F(x) = \big(f(x_1,x_2),\ g(x_1,x_2),\ h(x_3)\big)$, $x = (x_1,x_2,x_3)$, $x \in \mathbb{R}^3 \setminus \{(0,0,t) : t \in \mathbb{R}\}$, où f, g sont de classe C^1 dans $\mathbb{R}^2 \setminus \{(0,0)\}$ et $h \in C^1(\mathbb{R})$. Vérifier que :

(i) F est un champ irrotationnel si et seulement si $\mathrm{D}_1 g = \mathrm{D}_2 f$ dans $\mathbb{R}^2 \setminus \{(0,0)\}$;
(ii) F est un champ conservatif si et seulement si le champ $G = (f,g)$ dans $\mathbb{R}^2 \setminus \{(0,0)\}$ dérive d'un potentiel.

14 Soit Ω une partie ouverte de $\mathbb{R}^3$; soit $f : \Omega \to \mathbb{R}$ une fonction telle que

$$u(x) = \int_\Omega \frac{f(y)}{\|y-x\|}\,\mathrm{d}y$$

existe pour tout $x \in \mathbb{R}^3$, u étant une fonction de classe C^1 partout ; nous supposons, de plus, que l'on peut obtenir les dérivées partielles de u en dérivant sous le signe intégrale qui le définit. Vérifier que

$$F(x) = \operatorname{grad} u(x) = \int_\Omega \frac{f(y)(y-x)}{\|y-x\|^3}\,\mathrm{d}y\,, \quad x \in \mathbb{R}^3\ .$$

(Le champ F donné par $\operatorname{grad} u$ s'appelle le *champ newtonien* défini par f, u étant le potentiel (newtonien) pour F.)

15 On maintient les notations de l'exercice 14 et on suppose de plus que Ω est borné et que f est bornée et de classe C^1 dans Ω. Vérifier (en utilisant § 6.6.6) que

$$\operatorname{div} F(x) = \begin{cases} -4\pi f(x) & \text{si } x \in \Omega \\ 0 & \text{si } x \in \operatorname{ext} \Omega. \end{cases}$$

Si, de plus, f est telle que $u \in C^2(\mathbb{R}^3)$, alors $\operatorname{rot} F = 0$ partout.

REMARQUE. Les conditions sur f telles que u et donc F possèdent les différentes propriétés mentionnées dans les exercices 14 et 15 ont été étudiées très finement ; l'objectif de ces deux exercices est de montrer que, en général, les champs newtoniens sont irrotationnels et compressibles, les points x tels que $f(x) \neq 0$ étant les *sources* du champ F.

16 Soit Ω une partie ouverte de $\mathbb{R}^3$; soit $a : \Omega \to \mathbb{R}^3$ un champ vectoriel tel que

$$A(x) = \int_\Omega \frac{a(y)}{\|y - x\|}\, dy$$

existe pour tout $x \in \mathbb{R}^3$, A étant de classe C^1 partout ; nous supposons, de plus, que l'on peut obtenir les dérivées partielles de A en dérivant sous le signe intégrale qui le définit. Vérifier que

$$G(x) = \operatorname{rot} A(x) = \int_\Omega \frac{a(y) \times (x - y)}{\|x - y\|^3}\, dy\,, \quad x \in \mathbb{R}^3\,.$$

(Le champ G donné par $\operatorname{rot} A$ s'appelle le *champ de Biot et Savart* défini par a, A étant le potentiel vecteur pour G.)

17 On maintient les notations de l'exercice 16 et on suppose de plus que : Ω est borné, a est borné et de classe C^1 dans Ω. Vérifier (en utilisant § 6.6.6) que

$$\triangle A(x) = \begin{cases} -4\pi a(x) & \text{si } x \in \Omega \\ 0 & \text{si } x \in \operatorname{ext} \Omega. \end{cases}$$

Supposons que a est tel que $\operatorname{div} A(x) = 0$, $x \in \mathbb{R}^3$ et que A est de classe C^2 partout ; alors le champ G est tel que $\operatorname{div} G = 0$ partout et $\operatorname{rot} G(x) = 4\pi a(x)$ si $x \in \Omega$ et $\operatorname{rot} G(x) = 0$ si $x \in \operatorname{ext} \Omega$.

REMARQUE. On peut démontrer que $\operatorname{div} A = 0$ si $\operatorname{div} a = 0$ dans Ω et si Ω est un domaine régulier tel que $\nu(x) \cdot a(x) = 0$ pour $x \in \operatorname{Fr} \Omega$, $\nu(x)$ étant la normale unitaire au point x de $\operatorname{Fr} \Omega$; cette condition possède une bonne interprétation physique en électromagnétisme (cf. Chilov [9, p. 339] pour la preuve analytique).

18 Soient $f \in C^1(\Omega)$, $R \in C^1(\Omega;\mathbb{R}^3)$, $\operatorname{div} R = 0$, Ω étant une partie ouverte de $\mathbb{R}^3$. Un problème général de l'analyse vectorielle est de déterminer un champ F dans Ω tel que $\operatorname{div} F = f$, $\operatorname{rot} F = R$ dans Ω. Démontrer que ce problème possède une solution locale dans un voisinage ouvert de chaque point de Ω en procédant comme suit : soit $\xi \in \Omega$ et V un pavé ouvert centré à ξ tel que $V \subset \Omega$; puisque V est étoilé, il existe un champ G tel que $\operatorname{rot} G = R$ dans V ; si $\varphi = \operatorname{div} G$, prendre $F = G + \operatorname{grad} h$ où h est tel que $\Delta h = f - \varphi$ dans V ; alors $\operatorname{div} F = f$; $\operatorname{rot} F = R$ dans V.

6.7 EQUATIONS DE MAXWELL

6.7.1 Introduction

Cette section doit être considérée comme une prolongation des exercices du paragraphe 6.6.7, présentée dans la notation et la terminologie établies dans la théorie d'électromagnétisme ; une grande partie de cette dernière théorie est résumée dans les ***équations de Maxwell*** qui concernent quatre champs vectoriels E, D, H, B « inconnus » et une fonction scalaire ρ ainsi qu'un champ vectoriel J « connus » ; toutes ces quantités dépendent du « temps » t et de la « position » $x = (x_1, x_2, x_3)$; ainsi

$$E = E(t,x) = E(t,x_1,x_2,x_3)\,, \quad E = (E_1,E_2,E_3)\,,$$

et nous utilisons une écriture similaire pour les autres quantités D, H, B, ρ, J.

Les équations de Maxwell sont les suivantes :

$$\begin{aligned}
&\text{(i)}\ \operatorname{div} D = \rho && \text{(loi de Gauss-électrique)}\\
&\text{(ii)}\ \operatorname{div} B = 0 && \text{(loi de Gauss-magnétique)}\\
&\text{(iii)}\ \operatorname{rot} E = -\frac{\partial B}{\partial t} && \text{(loi de Maxwell-Faraday)}\\
&\text{(iv)}\ \operatorname{rot} H = J + \frac{\partial D}{\partial t} && \text{(loi de Maxwell-Ampère).}
\end{aligned} \tag{6.38}$$

E : ***champ électrique*** ; D : ***induction électrique*** ; H : ***champ magnétique*** ; B : ***induction magnétique*** ; ρ : ***densité de charge*** ; J : ***densité de courant***.

Le champ D se déduit de E et le champ B se déduit de H moyennant des ***relations constitutives*** du type $D = \varepsilon E$, $B = \mu H$; dans la théorie générale ε et μ sont des opérateurs linéaires de $\mathbb{R}^3$ dans $\mathbb{R}^3$ (appelés aussi *tenseurs*) dépendants de t et x ; ces opérateurs ε, μ sont déterminés par les considérations physiques du « milieu » ou « matériau » concerné ; si le milieu est « homogène et isotrope », on considère ε et μ comme des constantes scalaires (ε étant appelée la « constante diélectrique » et μ la « perméabilité magnétique »). Dans cette

section, nous ne tenons aucun compte ni de de cette terminologie ni de ce contexte physique ; ils sont évoqués pour des raisons « culturelles ». En fait, le contexte physique est bien plus compliqué par les discussions essentielles concernant le choix d'unités de mesure que nous avons ignoré complètement ; ceci fait que dans nos équations (6.38) on n'a pas les facteurs 4π ou $1/c$ (c étant la vitesse de la lumière) que l'on trouve dans ces formules dans la littérature physique. Le champ $D - E = P$ est appelé *vecteur de polarisation* et le champ $B - H$ est appelé *vecteur d'aimantation* (ou *magnétisation*).

Dans la suite, on supposera que les quantités E, D, etc. sont au moins deux fois continûment dérivables ; cela veut dire physiquement que l'on travaille dans les parties « vides » d'espace ou dans les parties où les champs sont distribués de manière lisse ; aussi, les opérations faites de dérivations sous le signe intégrale sont justifiées.

6.7.2 Interdépendance des équations (6.38)

De (i) et (iv) de (6.38), on obtient :

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \frac{\partial D}{\partial t} = \operatorname{div}(\operatorname{rot} H - J) = -\operatorname{div} J\,,$$

ce qui donne

$$\frac{\partial \rho}{\partial t} + \operatorname{div} J = 0\,, \tag{6.39}$$

relation dite de ***continuité*** ou de ***conservation de l'électricité***. En appliquant div sur les deux côtés de (iii) et (iv) de (6.38), on obtient :

$$0 = \operatorname{div} \operatorname{rot} E = -\frac{\partial}{\partial t}(\operatorname{div} B)$$

$$0 = \operatorname{div} \operatorname{rot} H = \operatorname{div} J + \frac{\partial}{\partial t}(\operatorname{div} D) = \frac{\partial}{\partial t}(\operatorname{div} D - \rho)\,,$$

compte tenu de (6.39) ; donc

$$\frac{\partial}{\partial t}(\operatorname{div} B) = 0\,, \qquad \frac{\partial}{\partial t}(\operatorname{div} D - \rho) = 0\,. \tag{6.40}$$

On conclut de (6.40) que si (i) et (ii) de (6.38) sont vérifiées à un instant t_0, alors elles sont vérifiées pour tout t ; donc les quatres équations de (6.38) peuvent être remplacées essentiellement par les deux équations (iii) et (iv) de (6.38) en ajoutant à celles-ci la relation de continuité (6.39).

6.7.3 Lois de conservation

La relation de continuité peut être écrite comme suit : soit Ω un domaine régulier dans $\mathbb{R}^3$, ν étant la normale unitaire dirigée vers l'extérieur définie sur $\operatorname{Fr} \Omega$; alors on a

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \rho \,\mathrm{d}x\right) = -\int_{\operatorname{Fr}\Omega} J \cdot \nu \,\mathrm{d}\sigma \tag{6.41}$$

car le côté gauche de (6.41) vaut

$$\int_\Omega \frac{\partial \rho}{\partial t}\,\mathrm{d}x = -\int_\Omega \operatorname{div} J\,\mathrm{d}x = -\int_{\mathrm{Fr}\,\Omega} J\cdot\nu\,\mathrm{d}\sigma$$

grâce au théorème de la divergence et à (6.39). La relation (6.41) est interprétée comme une loi de conservation de l'électricité. Les lois de conservation suivantes sont également immédiates à partir de (6.38)(i), (ii), (iii), (iv) respectivement :

$$\int_\Omega \rho\,\mathrm{d}x = \int_{\mathrm{Fr}\,\Omega} D\cdot\nu\,\mathrm{d}\sigma \tag{6.42}$$

$$\int_{\mathrm{Fr}\,\Omega} B\cdot\nu\,\mathrm{d}\sigma = 0 \tag{6.43}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega B\,\mathrm{d}x\right) = \int_{\mathrm{Fr}\,\Omega}(E\times\nu)\,\mathrm{d}\sigma \tag{6.44}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega D\,\mathrm{d}x\right) = -\int_\Omega J\,\mathrm{d}x - \int_{\mathrm{Fr}\,\Omega}(H\times\nu)\,\mathrm{d}\sigma\,. \tag{6.45}$$

Dans (6.44) et (6.45), on a utilisé le théorème de la divergence dans la forme de l'exercice 6(ii) (§ 6.5.7) :

$$\int_\Omega \operatorname{rot} F\,\mathrm{d}x = -\int_{\mathrm{Fr}\,\Omega}(F\times\nu)\,\mathrm{d}\sigma\,.$$

6.7.4 Circulations de E et H

En utilisant le théorème de Stokes, on peut reformuler (iii) et (iv) de (6.38) comme des relations concernant les circulations de E et H respectivement :

$$\int_{\partial D} E\cdot\mathrm{d}\vec{s} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_D B\cdot\nu\,\mathrm{d}\sigma\right) \tag{6.46}$$

$$\int_{\partial D} H\cdot\mathrm{d}\vec{s} = \int_D\left(J+\frac{\partial D}{\partial t}\right)\cdot\nu\,\mathrm{d}\sigma \tag{6.47}$$

où D est un domaine régulier dans une surface régulière orientée. On interprète (6.46) comme la loi de Faraday et (6.47) comme la loi d'Ampère. En fait, (6.46) équivaut (iii) de (6.38) et (6.47) équivaut (iv) de (6.38).

6.7.5 Vecteur de Poynting

Le champ vectoriel

$$S = E\times H$$

s'appelle le ***vecteur de Poynting***. On vérifie immédiatement, en utilisant la formule (7) du paragraphe 6.6.1 et les relations (iii) et (iv) de (6.38), que

$$\operatorname{div} S = -E\cdot J - \left(H\cdot\frac{\partial B}{\partial t} + E\cdot\frac{\partial D}{\partial t}\right)\,. \tag{6.48}$$

La relation (6.48) est interprétée comme « le bilan (local) de l'énergie électromagnétique ». Lorsque $B = H$, $D = E$ (« milieu vide »), (6.48) s'écrit comme

$$\operatorname{div} S = -E \cdot J - \frac{\partial \mathcal{E}}{\partial t} \tag{6.49}$$

où

$$\mathcal{E} = \frac{1}{2}\left(\|H\|^2 + \|E\|^2\right)$$

est la « densité d'énergie électromagnétique » relative au vide. En utilisant le théorème de la divergence, on peut écrire (6.49) comme suit :

$$\int_{\operatorname{Fr}\Omega} S \cdot \nu \, \mathrm{d}\sigma = -\int_{\Omega} E \cdot J \, \mathrm{d}x - \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{\Omega} \mathcal{E} \, \mathrm{d}x\right) \tag{6.50}$$

pour tout domaine régulier Ω (« dans le vide »), ν étant la normale unitaire dirigée vers l'extérieur associée à $\operatorname{Fr}\Omega$.

6.7.6 Solutions des équations

Supposons que $B = H$, $D = E$ dans (6.38) (« milieu vide ») ; les équations (6.38) deviennent alors

$$\begin{aligned} &\text{(i)}\ \operatorname{div} E = \rho\,, && \text{(ii)}\ \operatorname{div} H = 0\,, \\ &\text{(iii)}\ \operatorname{rot} E = -\frac{\partial H}{\partial t}\,, && \text{(iv)}\ \operatorname{rot} H = J + \frac{\partial E}{\partial t}\,. \end{aligned} \tag{6.51}$$

Etant donné ρ et J (satisfaisant la relation de continuité (6.39)), on peut résoudre (6.51) pour obtenir E et H comme suit. Nous supposons que E, H, ρ, J sont de classe C^2 au moins pour x dans un domaine Ω et $t \in \mathbb{R}$.

Notons d'abord que (E, H) *résout (ii) et (iii) de (6.51) dans un domaine étoilé* $\Omega \subset \mathbb{R}^3$ *et* $t \in \mathbb{R}$ *si et seulement s'il existe un potentiel scalaire* φ *et un potentiel vecteur* A *(définis dans* $\mathbb{R} \times \Omega$*) tels que*

$$H = \operatorname{rot} A\,, \quad E = -\frac{\partial A}{\partial t} - \operatorname{grad} \varphi\,. \tag{6.52}$$

En effet, si (E, H) satisfait (ii) et (iii) de (6.51) dans $\mathbb{R} \times \Omega$, Ω étoilé, alors H est de la forme $\operatorname{rot} A$ (à cause de (ii)) et (iii) entraîne

$$\operatorname{rot}\left(E + \frac{\partial A}{\partial t}\right) = 0$$

d'où le fait que $E + \partial A/\partial t$ soit de la forme $-\operatorname{grad}\varphi$ (car Ω est étoilé), le choix du signe $-$ étant purement traditionnel.

Réciproquement, si (E, H) sont de la forme (6.52), on vérifie directement (ii) et (iii) de (6.51) (même si Ω n'est pas étoilé). ♦

Si E et H sont donnés par (6.52), on vérifie que

$$\operatorname{div} E = \Box\varphi - \frac{\partial}{\partial t}\left(\operatorname{div} A + \frac{\partial\varphi}{\partial t}\right) \tag{6.53}$$

$$\operatorname{rot} H - \frac{\partial E}{\partial t} = \Box A + \operatorname{grad}\left(\operatorname{div} A + \frac{\partial\varphi}{\partial t}\right) \tag{6.54}$$

où

$$\Box = \frac{\partial^2}{\partial t^2} - \Delta = \frac{\partial^2}{\partial t^2} - (D_{11} + D_{22} + D_{33})$$

est l'opérateur d'alembertien, Δ opérant sur les variables dites spatiales x_1, x_2, x_3 et si $A = (A_1, A_2, A_3)$,

$$\Box A = (\Box A_1, \Box A_2, \Box A_3)\,.$$

On déduit de (6.53) et (6.54) que (E, H) *donné par le couple* (A, φ) *satisfait les quatre équations de (6.51) dans* $\mathbb{R} \times \Omega$ *si l'on a dans* $\mathbb{R} \times \Omega$

$$\Box A = J\,, \quad \Box\varphi = \rho\,, \quad \operatorname{div} A + \frac{\partial\varphi}{\partial t} = 0\,. \tag{6.55}$$

Une équation du type $\Box A = J$ s'appelle ***équation des ondes*** (***non homogène*** si $J \neq 0$, ***homogène*** si $J = 0$) ; c'est une équation dite du type *hyperbolique* et nous étudierons les solutions de ce type d'équations dans le volume 3.

Notons ici que les solutions de ces équations ne sont déterminées que si l'on donne $A(t_0, x)$, $\partial A/\partial t(t_0; x)$, $x \in \Omega$, à un point $t = t_0$ (*conditions initiales*) et si l'on fixe les *conditions au bord* (par exemple, les valeurs de $A(t, x)$ pour $x \in \operatorname{Fr}\Omega$). Si les deux premières équations de (6.55) peuvent être résolues une après l'autre, il faut encore voir que les solutions A, φ des deux premières équations satisfont la troisième équation de (6.55) que l'on appelle ***condition de Lorentz***.

Remarquons que si $a : \mathbb{R}^3 \to \mathbb{R}$ est une fonction de classe C^2, alors la transformation

$$(A, \varphi)_a = (\widetilde{A}, \widetilde{\varphi})\,, \quad \widetilde{A} = A + \operatorname{grad} a\,, \quad \widetilde{\varphi} = \varphi - \frac{\partial a}{\partial t} \tag{6.56}$$

(appelée ***transformation de jauge***) est telle que *si* (E, H) *est donné par (6.52) via* (A, φ), *alors on obtient le même couple* (E, H) *dans (6.52) via* $(\widetilde{A}, \widetilde{\varphi})$. La vérification simple de ceci est proposée comme un exercice ; *on vérifie de plus que*

$$\operatorname{div}\widetilde{A} + \frac{\partial\widetilde{\varphi}}{\partial t} = \operatorname{div} A + \frac{\partial\varphi}{\partial t} - \Box a\,. \tag{6.57}$$

Ceci permet de satisfaire la troisième équation de (6.55) par un choix approprié de a.

On en conclut aussi que *si* $\Box a = 0$, *alors si* (A, φ) *satisfait (6.55),* $(\widetilde{A}, \widetilde{\varphi}) = (A, \varphi)_a$ *satisfait aussi (6.55).* Ainsi, on voit que (6.55) ne détermine pas les deux potentiels A et φ uniquement.

L'existence et l'unicité des solutions de (6.51) dans un domaine $\Omega \subset \mathbb{R}^3$ et $t \in \mathbb{R}$, lorsque ρ, J (satisfaisant la relation de continuité (6.39)) sont donnés, est un des problèmes importants associés aux équations de Maxwell ; mais il y a un nombre incalculable d'autres problèmes associés fournis par la physique et la technique ; voir Dautray et Lions [12, vol. 2 (chap. 9)] pour plusieurs résultats mathématiques précis. Notre discussion dans cette section est basée sur [12, vol. 1, (chap. 1, sect. 4)] où l'on trouvera d'autres références importantes sur les équations de Maxwell.)

6.8 COORDONNÉES CURVILIGNES DANS $\mathbb{R}^3$

6.8.1 Introduction

Soient $\widetilde{\Omega}$, Ω deux domaines de $\mathbb{R}^3$ et $\varphi : \widetilde{\Omega} \to \Omega$ un C^1-difféomorphisme ; notons les points de $\widetilde{\Omega}$ par $u = (u_1, u_2, u_3)$, ceux de Ω par $x = (x_1, x_2, x_3)$ et les composantes de φ par $(\varphi_1, \varphi_2, \varphi_3)$. Si $x = \varphi(u)$, on dit que (u_1, u_2, u_3) sont les ***coordonnées curvilignes*** du point x définies par φ ; puisque φ est une bijection, chaque point $x \in \Omega$ possède un unique système de coordonnées curvilignes (u_1, u_2, u_3).

Soit $\psi : \Omega \to \widetilde{\Omega}$ la fonction inverse de φ :

$$\psi\big(\varphi(u)\big) = u\,, \quad u \in \widetilde{\Omega}\,.$$

On sait que ψ (ainsi que φ) est de classe C^1 et, si $x = \varphi(u)$,

$$\psi'(x)\,\varphi'(u) = I$$

où I est l'application linéaire identité de $\mathbb{R}^3$ dans $\mathbb{R}^3$. En particulier

$$J(u) = \det \varphi'(u) = \frac{\partial(\varphi_1, \varphi_2, \varphi_3)}{\partial(u_1, u_2, u_3)}\,(u) \neq 0\,, \quad u \in \widetilde{\Omega}\,.$$

Puisque $\widetilde{\Omega}$ est connexe, le signe de $J(u)$ sera le même pour tout $u \in \widetilde{\Omega}$.

Rappelons que la matrice associée à $\varphi'(u)$ est

$$\big[\mathrm{D}_j \varphi_i(u)\big]\,, \quad 1 \leqslant i\,,\ j \leqslant 3\,;$$

la condition $J(u) \neq 0$ équivaut à l'indépendance linéaire des trois vecteurs $\mathrm{D}_j\varphi(u)$, $j = 1, 2, 3$, les trois colonnes de la matrice associée à $\varphi'(u)$. La condition $J(u) > 0$, $u \in \widetilde{\Omega}$, signifie que la base ordonnée de $\mathbb{R}^3$ donnée par

$$\big\{\mathrm{D}_1\varphi(u),\ \mathrm{D}_2\varphi(u),\ \mathrm{D}_3\varphi(u)\big\}$$

pour un $u \in \widetilde{\Omega}$ fixé, introduit la même orientation en $\mathbb{R}^3$ (appelée ***positive***) que celle donnée par la base canonique ordonnée $\{e_1, e_2, e_3\}$ où

$$e_1 = (1,0,0)\,, \quad e_2 = (0,1,0)\,, \quad e_3 = (0,0,1)\,.$$

Soit $F = (F_1, F_2, F_3)$ un champ vectoriel dans Ω. Puisque $\{D_j\varphi(u),\ j = 1, 2, 3\}$ est une base pour $\mathbb{R}^3$ (pour $u \in \widetilde{\Omega}$ fixé), on peut écrire $F(x)$ $(x = \varphi(u))$ comme une unique combinaison linéaire des $D_j\varphi(u)$. Le problème que nous considérons dans cette section est celui *de représenter* div F, rot F, grad f, Δf *(les deux derniers pour* $f : \Omega \to \mathbb{R}$*) en termes des coordonnées de* $F(x)$ *dans la base* $\{D_j\varphi(u),\ j = 1, 2, 3\}$ *et les opérations différentielles* $\partial/\partial u_i$, $i = 1, 2, 3$. Il est sous-entendu que F et f sont suffisamment dérivables et que div F, rot f, grad f et Δf sont définis avec les opérations $\partial/\partial x_i$, $i = 1, 2, 3$.

Nous présentons la solution du problème posé pour le cas particulier important où les coordonnées curvilignes sont ***orthogonales*** dans le sens que

$$D_i\varphi(u) \cdot D_j\varphi(u) = 0\,, \quad i \neq j\,,\ u \in \widetilde{\Omega}\,,$$

Posons

$$\alpha_j(u) = \|D_j\varphi(u)\|\,, \quad \widetilde{e}_j(u) = \frac{1}{\alpha_j(u)} D_j\varphi(u)\,, \quad j = 1, 2, 3\,;$$

dans le cas des coordonnées orthogonales,

$$\{\widetilde{e}_1(u), \widetilde{e}_2(u), \widetilde{e}_3(u)\}$$

donne une base orthonormale pour $\mathbb{R}^3$ $(u \in \widetilde{\Omega}$ fixé) et

$$\big|J(u)\big| = \alpha_1(u)\,\alpha_2(u)\,\alpha_3(u) = \big[D_1\varphi(u), D_2\varphi(u), D_3\varphi(u)\big]\,, \quad u \in \Omega\,,$$

le dernier terme étant le produit mixte des trois vecteurs concernés. Dans la suite, nous supprimerons l'argument u dans ce genre d'écriture, si cela ne cause pas d'ambiguïté.

6.8.2 Notations et conventions supplémentaires

Les notations et conventions du paragraphe 6.8.1 étant admises, nous supposons dorénavant que les coordonnées curvilignes sont *orthogonales* et *positives* dans le sens que $J(u) > 0$, $u \in \widetilde{\Omega}$; les remarques appropriées seront intercalées pour indiquer les changements qu'il faut faire si $J(u) < 0$, $u \in \widetilde{\Omega}$; elle concernent *exclusivement* le calcul de rot F.

Si $\xi \in \mathbb{R}^3$, nous notons par $\widetilde{\xi}_i$, $i = 1, 2, 3$, les coordonnées de ξ dans la base orthonormale $\{\widetilde{e}_1, \widetilde{e}_2, \widetilde{e}_3\}$; donc

$$\widetilde{\xi}_i = \xi \cdot \widetilde{e}_i\,,\ i = 1, 2, 3\,, \quad \xi = \sum_{i=1}^{3} \xi_i e_i = \sum_{i=1}^{3} \widetilde{\xi}_i \widetilde{e}_i\,.$$

Notons que nous avons supprimé l'argument $u \in \widetilde{\Omega}$ dans l'écriture $\widetilde{e}_i$; même si ξ est un vecteur constant, les $\widetilde{\xi}_i$ sont des fonctions de u.

Nous convenons, de plus, que, si $g : \Omega \to \mathbb{R}$, l'écriture $g(x)$ représentera également $g\big(\varphi(u)\big)$ (où $x = \varphi(u)$) ; ainsi, g sera tantôt une fonction de x_1, x_2,

x_3, tantôt une fonction de u_1, u_2, u_3 ; une convention similaire vaut aussi pour $G : \Omega \to \mathbb{R}^3$. Cet abus de notation est largement accepté dans ce contexte car il allège les formules concernées ; s'il y a une ambiguïté d'interprétation possible, nous expliciterons les choses en détail.

6.8.3 Proposition

Les notations et conventions des paragraphes 6.8.1 et 6.8.2 restent en vigueur. Soient $f : \Omega \to \mathbb{R}$ et $F : \Omega \to \mathbb{R}^3$ de classe C^1 (de classe C^2 dans la formule (d)). Alors :

(a) $\displaystyle (\operatorname{grad} f)_i^{\sim} = \frac{1}{\alpha_i}\frac{\partial f}{\partial u_i}, \quad i = 1, 2, 3 ;$

(b) $\displaystyle \operatorname{div} F = \frac{1}{J}\sum_{i=1}^{3}\frac{\partial}{\partial u_i}\left(\frac{J}{\alpha_i}\widetilde{F}_i\right) ;$

(c) $\displaystyle \sum_{i=1}^{3}(\operatorname{rot} F)_i^{\sim}\widetilde{e}_i = \frac{1}{J}\begin{vmatrix} \alpha_1\widetilde{e}_1 & \alpha_2\widetilde{e}_2 & \alpha_3\widetilde{e}_3 \\ \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_3} \\ \alpha_1\widetilde{F}_1 & \alpha_2\widetilde{F}_2 & \alpha_3\widetilde{F}_3 \end{vmatrix}$

par exemple,

$$(\operatorname{rot} F)_1^{\sim} = \frac{\alpha_1}{J}\left\{\frac{\partial}{\partial u_2}(\alpha_3\widetilde{F}_3) - \frac{\partial}{\partial u_3}(\alpha_2\widetilde{F}_2)\right\} .$$

(d) $\displaystyle \Delta f = \frac{1}{J}\sum_{i=1}^{3}\frac{\partial}{\partial u_i}\left(\frac{J}{\alpha_i^2}\frac{\partial f}{\partial u_i}\right) .$

REMARQUE. Si $J < 0$, les expressions pour $(\operatorname{rot} F)_1^{\sim}$ doivent être multipliées par -1.

DÉMONSTRATION

Nous supposerons dans cette démonstration que le difféomorphisme $\varphi : \widetilde{\Omega} \to \Omega$ sous-jacent est de classe C^2 ; une autre démonstration que nous mentionnerons dans la suite fonctionnera avec φ de classe C^1.

(a) On a

$$(\operatorname{grad} f)_1^{\sim} = (\operatorname{grad} f)\cdot\widetilde{e}_1 = \frac{1}{\alpha_1}\sum_{j=1}^{3}\frac{\partial f}{\partial x_j}\frac{\partial \varphi_j}{\partial u_1} = \frac{1}{\alpha_1}\frac{\partial f}{\partial u_1}$$

selon la formule de la dérivation des fonctions composées appliquées à

$$\frac{\partial}{\partial u_1} f\big(\varphi_1(u), \varphi_2(u), \varphi_3(u)\big) ;$$

les formules pour $(\operatorname{grad} f)_i^{\sim}$, pour $i = 2, 3$, sont obtenues par le même raisonnement.

Explicitons le sens de cette formule : le membre de gauche est évalué à un point $x = \varphi(u)$; le membre de droite est évalué au point u, f étant interprétée comme $f \circ \varphi$.

Avant de commencer avec les démonstrations de (b), (c) et (d), nous établissons quelques propriétés de $\tilde{e}_1$, $\tilde{e}_2$, $\tilde{e}_3$. Notons d'abord que (J étant > 0), on a

$$\tilde{e}_1 \times \tilde{e}_2 = \tilde{e}_3\,, \quad \tilde{e}_2 \times \tilde{e}_3 = \tilde{e}_1\,, \quad \tilde{e}_3 \times \tilde{e}_1 = \tilde{e}_2\,; \tag{6.58}$$

c'est une conséquence immédiate de l'orthonormalité des $\tilde{e}_j$ et du fait que le produit mixte

$$[\tilde{e}_1, \tilde{e}_2, \tilde{e}_3] = \frac{J}{\alpha_1\alpha_2\alpha_3} = 1\,.$$

Etablissons maintenant que

$$\operatorname{rot}\left(\frac{1}{\alpha_i}\,\tilde{e}_i\right) = 0\,, \quad i = 1, 2, 3 \tag{6.59}$$

et

$$\operatorname{div}\left(\frac{\alpha_i}{J}\,\tilde{e}_i\right) = 0\,, \quad i = 1, 2, 3\,. \tag{6.60}$$

En effet, si $\varphi^{-1} = \psi = (\psi_1, \psi_2, \psi_3)$, on a

$$\operatorname{grad}\psi_i(x) = \frac{1}{\alpha_i(u)}\,\tilde{e}_i(u)\,, \quad i = 1, 2, 3\,,\ x = \varphi(u)\,, \tag{6.61}$$

ce qui est une conséquence de

$$[\mathrm{D}_j\psi_i(x)]\ [\mathrm{D}_j\varphi_i(u)] = I\,;$$

ainsi, $\operatorname{grad}\psi_1$, la première ligne de la matrice $[\mathrm{D}_j\psi_i]$, est orthogonal aux vecteurs $\mathrm{D}_2\varphi$, $\mathrm{D}_3\varphi$, les deux dernières colonnes de $[\mathrm{D}_j\varphi_i]$ d'où

$$\operatorname{grad}\psi_1 = k\,\mathrm{D}_1\varphi$$

pour un k réel ; or, de la même relation matricielle ci-dessus, on obtient

$$\operatorname{grad}\psi_1 \cdot \mathrm{D}_1\varphi = 1$$

d'où $k\alpha_1^2 = 1$, ce qui donne (6.61) pour $i = 1$ car $\tilde{e}_1 = (\mathrm{D}_1\varphi)/\alpha_1$; pour $i = 2, 3$, le raisonnement est analogue.

La formule (6.59) est une conséquence immédiate de (6.61) car le rotationnel d'un gradient est toujours nul ; ici, nous utilisons le fait que les φ_i, donc les ψ_i, sont de classe C^2. De plus, de (6.61), on a

$$\operatorname{grad}\psi_1 \times \operatorname{grad}\psi_2 = \frac{1}{\alpha_1\alpha_2}\,(\tilde{e}_1 \times \tilde{e}_2) = \frac{1}{\alpha_1\alpha_2}\,\tilde{e}_3 = \frac{\alpha_3}{J}\,\tilde{e}_3$$

d'où (selon la formule (14), § 6.6.1) on obtient (6.60) pour $i = 3$; les cas $i = 1, 2$ de (6.60) sont établis de la même façon.

(b) On a

$$F = \sum_{i=1}^{3} \widetilde{F}_i \widetilde{e}_i = \sum_i \left(\frac{J}{\alpha_i}\widetilde{F}_i\right)\left(\frac{\alpha_i}{J}\widetilde{e}_i\right)$$

d'où

$$\begin{aligned}\operatorname{div} F &= \sum_i \operatorname{div}\left[\left(\frac{J}{\alpha_i}\widetilde{F}_i\right)\left(\frac{\alpha_i}{J}\widetilde{e}_i\right)\right] \\ &= \sum_i \operatorname{grad}\left(\frac{J}{\alpha_i}\widetilde{F}_i\right)\cdot\left\{\frac{\alpha_i}{J}\widetilde{e}_i\right\}\end{aligned}$$

en utilisant la formule (6) (§ 6.6.1) et (6.60) ; ainsi de (a), on obtient

$$\begin{aligned}\operatorname{div} F &= \sum_i \frac{\alpha_i}{J}\left\{\operatorname{grad}\left(\frac{J}{\alpha_i}\widetilde{F}_i\right)\right\}^{\sim}_i = \sum_i \frac{\alpha_i}{J}\cdot\frac{1}{\alpha_i}\frac{\partial}{\partial u_i}\left(\frac{J}{\alpha_i}\widetilde{F}_i\right) \\ &= \frac{1}{J}\sum_i \frac{\partial}{\partial u_i}\left(\frac{J}{\alpha_i}\widetilde{F}_i\right).\end{aligned}$$

(c) On a

$$F = \sum_{i=1}^{3} \widetilde{F}_i \widetilde{e}_i = \sum_i (\alpha_i \widetilde{F}_i)\left(\frac{1}{\alpha_i}\widetilde{e}_i\right)$$

d'où

$$\operatorname{rot} F = \sum_i \operatorname{rot}\left[(\alpha_i\widetilde{F}_i)\left(\frac{1}{\alpha_i}\widetilde{e}_i\right)\right] = \sum_i \operatorname{grad}(\alpha_i\widetilde{F}_i)\times\left(\frac{1}{\alpha_i}\widetilde{e}_i\right)$$

selon la formule (8) (§ 6.6.1) et (6.59). Maintenant, compte tenu de (6.58) on obtient

$$\begin{aligned}(\operatorname{rot} F)^{\sim}_1 &= (\operatorname{rot} F)\cdot\widetilde{e}_1 \\ &= \operatorname{grad}(\alpha_2\widetilde{F}_2)\times\left(\frac{1}{\alpha_2}\widetilde{e}_2\right)\cdot\widetilde{e}_1 + \operatorname{grad}(\alpha_3\widetilde{F}_3)\times\left(\frac{1}{\alpha_3}\widetilde{e}_3\right)\cdot\widetilde{e}_1 \\ &= \frac{1}{\alpha_2}\left\{\operatorname{grad}(\alpha_2\widetilde{F}_2)\right\}^{\sim}_3\left\{(\widetilde{e}_3\times\widetilde{e}_2)\cdot\widetilde{e}_1\right\} + \frac{1}{\alpha_3}\left\{\operatorname{grad}(\alpha_3\widetilde{F}_3)\right\}^{\sim}_2\left\{(\widetilde{e}_2\times\widetilde{e}_3)\cdot\widetilde{e}_1\right\} \\ &= \frac{1}{\alpha_2\alpha_3}\left\{\frac{\partial}{\partial u_2}(\alpha_3\widetilde{F}_3) - \frac{\partial}{\partial u_3}(\alpha_2\widetilde{F}_2)\right\}\end{aligned}$$

où l'on a utilisé (a). Les formules pour $(\operatorname{rot} F)^{\sim}_i$, $i = 2, 3$, sont obtenues par le même raisonnement.

(d) Ici f est de classe C^2 ; alors

$$\Delta f = \operatorname{div}(\operatorname{grad} f) = \frac{1}{J}\sum_{i=1}^{3}\frac{\partial}{\partial u_i}\left\{\frac{J}{\alpha_i}\,(\operatorname{grad} f)_i^{\sim}\right\}$$

$$= \frac{1}{J}\sum_{i=1}^{3}\frac{\partial}{\partial u_i}\left\{\frac{J}{\alpha_i^2}\,\frac{\partial f}{\partial u_i}\right\}$$

selon (b) et (a) successivement. ♦

REMARQUE. Si $J < 0$, le seul changement dans la démonstration précédente concerne (6.58) ; par exemple, si $J < 0$,

$$\widetilde{e}_1 \times \widetilde{e}_2 = -\widetilde{e}_3$$

car

$$[\widetilde{e}_1, \widetilde{e}_2, \widetilde{e}_3] = (\widetilde{e}_1 \times \widetilde{e}_2)\cdot \widetilde{e}_3 = \frac{J}{\alpha_1\alpha_2\alpha_3} = -1\,;$$

ceci fait que les expressions pour $(\operatorname{rot} F)_i^{\sim}$ dans (c) sont multipliées par -1.

6.8.4 Esquisse d'une autre démonstration

On a vu dans l'exercice 7 (§ 6.4.9) que $\operatorname{div} F$ peut être calculé comme

$$\operatorname{div} F(a) = \lim \frac{1}{\lambda_3(D)}\,\operatorname{Flux}(F; \operatorname{Fr} D)$$

où les domaines réguliers D convergent vers le point a dans un sens précis ; ceci montre immédiatement que $\operatorname{div} F$ ne dépend pas d'un choix particulier des coordonnées. Pour obtenir une expression analytique en termes des coordonnées curvilignes données par $\varphi : \widetilde{\Omega} \to \Omega$, on peut procéder comme suit :

Soit $a \in \Omega$, $v \in \widetilde{\Omega}$, $\varphi(v) = a$; soit $h = (h_1, h_2, h_3)$ (h_1, h_2, h_3 positifs, petits) ; posons

$$\widetilde{D} = \{u : v_i < u_i < v_i + h_i\,,\ i = 1, 2, 3\}$$

et $D = \varphi(\widetilde{D})$. On a $\operatorname{Fr} D = \varphi(\operatorname{Fr}\widetilde{D})$; $\operatorname{Fr} D$ se décompose en fermeture de six morceaux de surfaces, images par φ des six faces de $\widetilde{D}$. Prenons typiquement deux de ces morceaux :

$$S_1 = \varphi\big(\{u : u_1 = v_1\,,\ v_i < u_i < v_i + h_i\,,\ i = 2, 3\}\big)$$
$$S_2 = \varphi\big(\{u : u_1 = v_1 + h_1\,,\ v_i < u_i < v_i + h_i\,,\ i = 2, 3\}\big)\,.$$

On obtient que le flux de F à travers $S_1 \cup S_2$ est égal à

$$\mathrm{Flux}(F; S_2) + \mathrm{Flux}(F; S_1) = \iint_{\substack{u_1 = v_1 + h_1 \\ v_i < u_i < v_i + h_i \\ i=2,3}} [F \circ \varphi, \mathrm{D}_2\varphi_2, \mathrm{D}_3\varphi_3]\, du_2\, du_3$$

$$- \iint_{\substack{u_1 = v_1 \\ v_i < u_i < v_i + h_i \\ i=2,3}} [F \circ \varphi, \mathrm{D}_2\varphi_2, \mathrm{D}_3\varphi_3]\, du_2\, du_3$$

compte tenu des directions de la normale unitaire ν dirigée vers l'extérieur de D ; si h_1, h_2, h_3 sont petits, cette expression vaut approximativement

$$h_2 h_3 \left\{ [F \circ \varphi, \mathrm{D}_2\varphi_2, \mathrm{D}_3\varphi_3]_{v+h_1 e_1} - [F \circ \varphi, \mathrm{D}_2\varphi_2, \mathrm{D}_3\varphi_3]_v \right\}$$

$$= h_2 h_3 \left\{ \left(\frac{1}{\alpha_2 \alpha_3} \widetilde{F}_1 \right)_{v+h_1 e_1} - \left(\frac{1}{\alpha_2 \alpha_3} \widetilde{F}_1 \right)_v \right\}$$

$$= h_1 h_2 h_3 \frac{\partial}{\partial u_1} \left(\frac{1}{\alpha_2 \alpha_3} \widetilde{F}_1 \right) \Big|_v$$

car

$$[F \circ \varphi, \mathrm{D}_2\varphi_2, \mathrm{D}_3\varphi_3] = F \circ \varphi \cdot (\mathrm{D}_2\varphi_2 \times \mathrm{D}_3\varphi_3) = \frac{F \circ \varphi}{\alpha_2 \alpha_3} \cdot (\widetilde{e}_2 \times \widetilde{e}_3)$$

$$= \frac{F \circ \varphi \cdot \widetilde{e}_1}{\alpha_2 \alpha_3} = \frac{\widetilde{F}_1}{\alpha_2 \alpha_3}$$

où on a supposé que $J > 0$; si $J < 0$, il y aura un signe (-1) ici de sorte que, au total

$$\mathrm{Flux}(F; S_1 \cup S_2) \doteq h_1 h_2 h_3 \frac{\partial}{\partial u_1} \left(\frac{1}{\alpha_2 \alpha_3} \widetilde{F}_1 \right) \Big|_v .$$

Puisque $\lambda_3(D) \doteq (h_1 h_2 h_3) J(v)$, on obtient

$$\lim \frac{1}{\lambda_3(D)} \mathrm{Flux}(F; S_1 \cup S_2) = \frac{1}{J} \frac{\partial}{\partial u_1} \left(\frac{1}{\alpha_2 \alpha_3} \widetilde{F}_1 \right)$$

où la limite est prise lorsque h_1, h_2, $h_3 \to 0$; cette égalité précise aussi l'écriture $\doteq$; l'expression du côté droit ci-dessus est calculée au point v où $\varphi(v) = a$. On voit ainsi comment obtenir la formule (b) de la proposition 6.8.3 par ce raisonnement.

La formule (d) de la proposition 6.8.3 est une conséquence immédiate de (a) et (b) comme l'on a déjà vu ; la formule (c) de la proposition 6.8.3 concernant le calcul de $\mathrm{rot}\, F$ peut être faite en utilisant l'exercice 7 ou 8 (§ 6.5.7) :

$$\mathrm{rot}\, F(a) = \lim \frac{1}{\lambda_3(D)} \int_{\mathrm{Fr}\, D} (\nu \times F)\, d\sigma$$

$$\nu \cdot \mathrm{rot}\, F(a) = \lim \frac{1}{\beta_n} \int_C F \cdot d\vec{s}$$

où, dans le premier, D est une famille de domaines réguliers qui converge vers le point a et, dans le second, C est une famille de courbes orientées dans une surface contenant a qui converge vers le point a dans un sens évident. Les raisonnements à utiliser sont ceux indiqués ci-dessus pour le cas de la divergence ; nous laissons les détails en exercice (non banal).

6.8.5 Exercices

[1] Soit $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ une application linéaire orthogonale de déterminant 1. Vérifier que, dans ce cas, les expressions grad, div, rot, $\triangle$ restent invariantes.

[2] (Coordonnées cylindriques) On écrit :

$$x_1 = r\cos\theta\,, \quad x_2 = r\sin\theta\,, \quad x_3 = z\,;$$

ici $u_1 = r$, $u_2 = \theta$, $u_3 = z$. Vérifier qu'il s'agit des coordonnées orthogonales avec $\alpha_1 = 1$, $\alpha_2 = r$, $\alpha_3 = 1$, $J = r$ et

$$\operatorname{div} F = \frac{1}{r}\frac{\partial}{\partial r}(r\widetilde{F}_r) + \frac{1}{r}\frac{\partial}{\partial\theta}(\widetilde{F}_\theta) + \frac{\partial}{\partial z}(\widetilde{F}_z)$$

$$\operatorname{rot} F = \begin{vmatrix} \frac{1}{r}\widetilde{e}_r & \widetilde{e}_\theta & (1/r)\widetilde{e}_z \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial\theta} & \frac{\partial}{\partial z} \\ \widetilde{F}_r & r\widetilde{F}_\theta & \widetilde{F}_z \end{vmatrix}$$

$$\triangle f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial\theta^2} + \frac{\partial^2 f}{\partial z^2}$$

où $\widetilde{F}_r = \widetilde{F}_1$, $\widetilde{e}_r = \widetilde{e}_1$, etc.

[3] (Coordonnées sphériques) On écrit :

$$x_1 = r\sin\theta\cos\varphi\,, \quad x_2 = r\sin\theta\sin\varphi\,, \quad x_3 = r\cos\theta\,;$$

ici $u_1 = r$, $u_2 = \theta$, $u_3 = \varphi$. Vérifier qu'il s'agit des coordonnées orthogonales avec $\alpha_1 = 1$, $\alpha_2 = r$, $\alpha_3 = r\sin\theta$, $J = r^2\sin\theta$ et

$$\operatorname{div} F = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2\widetilde{F}_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\widetilde{F}_\theta) + \frac{1}{r\sin\theta}\frac{\partial\widetilde{F}_\varphi}{\partial\varphi}$$

$$\operatorname{rot} F = \begin{vmatrix} \frac{1}{r^2\sin\theta}\widetilde{e}_r & \frac{1}{r\sin\theta}\widetilde{e}_\theta & \frac{1}{r}\widetilde{e}_\varphi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial\theta} & \frac{\partial}{\partial\varphi} \\ \widetilde{F}_r & r\widetilde{F}_\theta & r\sin\theta\widetilde{F}_\varphi \end{vmatrix}$$

$$\triangle f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\varphi^2}$$

où $\widetilde{F}_r = \widetilde{F}_1$, $\widetilde{e}_r = \widetilde{e}_1$, etc.

REMARQUES. Il existe des tables qui donnent des formules pour une dizaine d'autres coordonnées curvilignes orthogonales importantes pour différentes applications (Spiegel [40] ou Madelung [25] ; ce dernier contient beaucoup d'autres formules utiles de l'analyse vectorielle). On peut développer aussi les formules pour les coordonnées curvilignes non orthogonales (en utilisant un formalisme appelé « calcul tensoriel ») mais dans les applications techniques, ce sont les coordonnées orthogonales qui semblent être les plus utiles.

6.9 COMMENTAIRES

Les appelations, théorème de Green, théorème de Gauss, théorème de Stokes sont purement conventionnelles. Le théorème de Stokes est clairement dû à Thomson (plus connu sous le nom de Lord Kelvin) qui signala ce résultat à Stokes dans une lettre datée de 1850 ; il apparaît comme un des problèmes de *Smith's Prize Examination* (de 1854), un examen annuel de l'Université de Cambridge dont le responsable était Stokes pendant les années 1849 à 1882. Mais il faut comprendre que ce théorème, comme les autres, avait des antécédents plus ou moins proches. Ainsi, le théorème de Green (datant de 1828 approximativement, Green ne se contenta pas de l'intégration dans le plan) est un précurseur évident du théorème de Stokes et celui-ci se trouve déjà dans les écrits d'Euler de 1771-1772 ; Riemann, dans sa thèse de 1851, donne une version du théorème de Green valable sur une surface (dite « surface de Riemann » aujourd'hui) sans mentionner ni Green ni aucun autre auteur dans ce contexte. On peut aussi associer le nom d'Ampère qui avait utilisé un cas particulier du théorème de Stokes dans ses travaux d'électromagnétisme.

Ce que nous avons appelé « les formules de Green » sont aussi nommées quelquefois comme « le théorème de Green ». Green ne semble pas connaître le théorème de la divergence de Gauss alors qu'une de ces formules est presque équivalente à celui-ci. Le théorème de la divergence est souvent attribué à Gauss qui avait utilisé ce théorème dans un cas particulier important dès 1813 ; mais sa forme générale est due à Ostrogradski (en 1828 dans $\mathbb{R}^3$, en 1834 dans $\mathbb{R}^N$, cf. Chilov [9, p. 357]). Par ailleurs, Lagrange avait déjà utilisé une forme de ce théorème avant Gauss et l'expression de divergence dans sa forme

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$$

apparaît de nombreuses fois dans les travaux antérieurs d'Euler.

Le langage de l'analyse vectorielle utilisé dans ce chapitre est dû, en grande partie, à Gibbs (1881-1884, dans des écrits de diffusion limitée) ; ce langage trouva un grand nombre d'adhérents avec la publication de son ouvrage *Vector Analysis* en 1901 (avec Wilson). Maxwell utilise le mot « rotation » (pour notre « rotationnel » ; le mot anglais standard aujourd'hui pour ceci est « *curl* »

aussi dû, en partie, à Maxwell) et « convergence » (pour la négative de notre « divergence », introduit aussi par Maxwell) mais il ne se sert pas des notations vectorielles dans son grand traité *Treatise on Electricity and Magnetism* de 1873, écrit exclusivement avec les équations cartésiennes dans les variables comme x, y, z. Mais déjà à partir de 1890 les exposés importants de la théorie de Maxwell sont présentés dans le langage d'analyse vectorielle par Heaviside et Föppl. Depuis ce temps jusqu'à nos jours, c'est le langage d'analyse vectorielle qui reste celui le plus utilisé pour la description et l'étude des phénomènes physiques.

Les démonstrations données pour les théorèmes principaux (théorèmes de Green, Gauss, Stokes) par les pionniers sont très sommaires du point de vue des mathématiques d'aujourd'hui. Les notions géométriques comme celles de surface, d'orientation, etc., et les notions exactes d'analyse comme celles de continuité, intégrabilité, etc. sont à peine abordées et les démonstrations présentées au début ne sont rien de plus que les vérifications que nous avons données dans ce chapitre. Il faut admettre que les difficultés inhérentes à la formulation exacte et générale de ces théorèmes et à leurs démonstrations sont considérables ; notre chapitre 7 est essentiellement dédié à l'accomplissement de ces tâches.

Il est intéressant de remarquer que l'année même de la publication de l'ouvrage de Gibbs-Wilson mentionné ci-dessus, apparaît un article important de Ricci et Levi-Civita (*Méthodes de calcul différentiel absolu et leurs applications*, *Mathematische Annalen*, vol. 54, 1901) qui introduit une analyse plus générale et plus profonde appelée aujourd'hui « analyse tensorielle ». Elle devient vite indispensable dans plusieurs domaines de mathématique et de physique (citons par exemple la géométrie différentielle, la relativité générale et la théorie d'élasticité) ; les développements du chapitre suivant sont une émanation de cette analyse tensorielle.

CHAPITRE 7

Théorème de Stokes généralisé

Dans ce chapitre nous donnons une version générale du théorème de Stokes dans $\mathbb{R}^N$ qui contient les théorèmes de Gauss, Green et Stokes dans $\mathbb{R}^2$ et $\mathbb{R}^3$ formulés dans le chapitre précédent. Pour plus de clarté, une version du théorème de la divergence précède le théorème de Stokes généralisé même si ce dernier la contient comme cas particulier. Plusieurs notions nécessaires pour une formulation précise et pour une démonstration complète du théorème de Stokes généralisé sont traitées en détail. Une discussion des formes fermées et exactes termine le chapitre.

7.1 FORMALISME DES FORMES DIFFÉRENTIELLES

7.1.1 Introduction

Dans cette section nous donnons les règles formelles d'un calcul symbolique qui permet d'englober tous les résultats d'analyse vectorielle vus précédemment et de les généraliser à $\mathbb{R}^N$. Pour $N = 2$ ou 3, on peut attribuer une interprétation immédiate à certaines expressions introduites ici mais l'objectif à présent est de nous familiariser avec leurs maniements algébriques ; ils sont très simples. N'essayez pas d'attacher, maintenant, un sens physique ou géométrique quelconque à ces opérations algébriques formelles ; la connaissance de l'analyse vectorielle présentée précédemment donne une première indication que ces manipulations symboliques peuvent être utiles.

Dans la suite de ce chapitre, nous reprendrons ces calculs au point de vue de l'algèbre multilinéaire. Les lecteurs qui ne veulent pas entrer dans ces détails peuvent se contenter de la lecture de la présente section seulement pour obtenir une idée claire de l'utilisation des formes différentielles pour la généralisation des théorèmes de Green, Gauss et Stokes aux dimensions supérieures. Ceux qui veulent connaître les démonstrations détaillées qui justifient ces théorèmes doivent naturellement poursuivre avec les sections suivantes.

Dans la suite de cette section, Ω est une partie ouverte non vide de $\mathbb{R}^N$.

7.1.2 Symbolisme des formes différentielles extérieures

Une ***forme différentielle extérieure de degré*** 0 dans Ω est une fonction réelle f définie dans Ω. Une ***forme différentielle extérieure de degré*** 1 dans Ω est une expression

$$\omega = f_1\,\mathrm{d}x_1 + \cdots + f_N\,\mathrm{d}x_N$$

où les f_i sont des fonctions réelles définies dans Ω ; une forme différentielle extérieure de degré 1 dans Ω s'appelle aussi une ***forme de Pfaff*** dans Ω. Une ***forme différentielle extérieure de degré*** 2 dans Ω est une expression du type

$$\omega = \sum_{1\leqslant i,\,j\leqslant N} f_{ij}\,\mathrm{d}x_i \wedge \mathrm{d}x_j$$

où les f_{ij} sont des fonctions réelles dans Ω. Plus généralement, une ***forme différentielle extérieure de degré*** p dans Ω est une expression de type

$$\omega = \sum_{1\leqslant i_1,\ldots,i_p\leqslant N} f_{i_1,\ldots,i_p}\,\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$$

où les $f_{i_1,\ldots,i_p}$ sont des fonctions réelles définies dans Ω ; nous appellerons ces fonctions les ***coefficients*** ou ***composantes*** de ω. Une forme différentielle extérieure de degré p dans Ω sera aussi appelée brièvement une ***p-forme différentielle*** ou encore une ***p-forme*** dans Ω, p étant un entier $\geqslant 0$.

Les symboles $\mathrm{d}x_i$ seront considérés comme des indéterminées soumises aux règles dites d'***anticommutation*** suivantes :

$$\mathrm{d}x_i \wedge \mathrm{d}x_i = 0\,, \quad \mathrm{d}x_i \wedge \mathrm{d}x_j = -\mathrm{d}x_j \wedge \mathrm{d}x_i \quad \text{si } i \neq j\ ;$$

de plus, ils satisfont la règle d'associativité :

$$(\mathrm{d}x_i \wedge \mathrm{d}x_j) \wedge \mathrm{d}x_k = \mathrm{d}x_i \wedge (\mathrm{d}x_j \wedge \mathrm{d}x_k)$$

de sorte qu'une expression comme

$$\mathrm{d}x_{i_1} \wedge \mathrm{d}x_{i_2} \wedge \cdots \wedge \mathrm{d}x_{i_p}$$

n'a pas besoin d'avoir de parenthèses. Les p-formes dans Ω sont donc des expressions polynomiales homogènes de degré p dans les indéterminées $\mathrm{d}x_1, \ldots, \mathrm{d}x_N$, ayant des coefficients qui sont les fonctions réelles définies dans Ω. En vue des règles d'anticommutation, les p-formes, pour $p > N$, sont nulles et une p-forme pour $p \leqslant N$ s'écrit uniquement comme

$$\omega = \sum_{1\leqslant i_1<i_2<\ldots<i_p\leqslant N} f_{i_1,\ldots,i_p}\,\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}\ ;$$

on dira que c'est l'***expression*** ou l'***écriture canonique*** pour ω. Deux p-formes sont dites ***égales*** si leurs composantes dans leurs expressions canoniques sont égales. Une p-forme qui a toutes ses composantes nulles sera notée 0 ; ainsi, 0 est une p-forme pour tout $p \geqslant 0$.

A titre d'exemple, une N-forme possède l'écriture canonique

$$f\,\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_N\ ;$$

une 2-forme en trois variables ($N = 3$) a l'écriture canonique

$$a\,\mathrm{d}x_1 \wedge \mathrm{d}x_2 + b\,\mathrm{d}x_1 \wedge \mathrm{d}x_3 + c\,\mathrm{d}x_2 \wedge \mathrm{d}x_3$$

ce qui est égal à

$$c\,\mathrm{d}x_2 \wedge \mathrm{d}x_3 - b\,\mathrm{d}x_3 \wedge \mathrm{d}x_1 + a\,\mathrm{d}x_1 \wedge \mathrm{d}x_2$$

Les p-formes constituent un espace vectoriel réel de dimension $\binom{N}{p}$; on les ajoute composante par composante et la multiplication par un nombre réel (ou par une fonction réelle) se fait, de même, composante par composante. Le ***produit extérieur*** d'une p-forme α et une q-forme β est une $(p+q)$-forme, notée $\alpha \wedge \beta$, formée en multipliant, d'une manière polynomiale, une expression pour α par une expression pour β, en tenant compte des règles de distributivité, associativité et anticommutation des symboles $\mathrm{d}x_i$. Résumons les propriétés du produit extérieur comme suit. On a

$$(\alpha + \beta) \wedge \gamma = (\alpha \wedge \gamma) + (\beta \wedge \gamma)$$

si α, β sont deux p-formes et γ est une q-forme ;

$$\alpha \wedge (\beta + \gamma) = (\alpha \wedge \beta) + (\alpha \wedge \gamma)$$

si α est une p-forme et β, γ sont deux q-formes ;

$$t\alpha \wedge \beta = \alpha \wedge t\beta = t(\alpha \wedge \beta)$$

si $t \in \mathbb{R}$ et α est une p-forme, β est une q-forme ; nous appellerons ces propriétés celles de ***distributivité*** du produit extérieur. La propriété d'***associativité*** est la suivante :

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma)$$

où α est une p-forme, β une q-forme, γ une r-forme ; grâce à cette propriété, on peut écrire $\alpha \wedge \beta \wedge \gamma$ ou, plus généralement, $\alpha_1 \wedge \cdots \wedge \alpha_n$; noter que si α_i est une p_i-forme, $\alpha_1 \wedge \cdots \wedge \alpha_n$ sera une $(p_1 + \cdots + p_n)$-forme. Finalement, on a la règle d'***anticommutation*** générale :

$$\alpha \wedge \beta = (-1)^{pq}(\beta \wedge \alpha)$$

si α est une p-forme et β est une q-forme. Toutes ces propriétés se vérifient facilement à partir des règles de manipulations des symboles $\mathrm{d}x_i$ étendues aux monômes

$$a\,\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$$

A titre d'exemple, si $N = 2$,

$$(f_1\,dx_1 + f_2\,dx_2) \wedge (g_1\,dx_1 + g_2\,dx_2) = (f_1g_2 - f_2g_1)\,dx_1 \wedge dx_2\ ;$$

noter que

$$(f_1g_2 - f_2g_1) = \det \begin{bmatrix} f_1 & f_2 \\ g_1 & g_2 \end{bmatrix} = -\det \begin{bmatrix} g_1 & g_2 \\ f_1 & f_2 \end{bmatrix}$$

ce qui correspond à la relation

$$(g_1\,dx_1 + g_2\,dx_2) \wedge (f_1\,dx_1 + f_2\,dx_2) = -(f_1g_2 - f_2g_1)\,dx_1 \wedge dx_2\,.$$

Soit $N = 3$; posons

$$\alpha = A_1\,dx_1 + A_2\,dx_2 + A_3\,dx_3\,, \quad \beta = B_1\,dx_1 + B_2\,dx_2 + B_3\,dx_3\ ;$$

on obtient

$$\begin{aligned}\alpha \wedge \beta = {} & (A_2B_3 - A_3B_2)\,dx_2 \wedge dx_3 + (A_3B_1 - A_1B_3)\,dx_3 \wedge dx_1 \\ & + (A_1B_2 - A_2B_1)\,dx_1 \wedge dx_2\ ;\end{aligned}$$

noter que les coefficients de $\alpha \wedge \beta$ sont formés de $A \times B$ où $A = (A_1, A_2, A_3)$, $B = (B_1, B_2, B_3)$ et on a bien

$$\alpha \wedge \beta = -\beta \wedge \alpha\,,$$

ce qui correspond à avoir $A \times B = -B \times A$. Si $\gamma = C_1\,dx_1 + C_2\,dx_2 + C_3\,dx_3$ on vérifie aussi que

$$\alpha \wedge \beta \wedge \gamma = \det \begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{bmatrix} dx_1 \wedge dx_2 \wedge dx_3\ ;$$

les propriétés des déterminants sont maintenant perçues comme celles des produits extérieurs. Plus généralement, si

$$\alpha_i = \sum_{j=1}^{N} a_{ij}\,dx_j\,, \quad i = 1, \ldots, N\,,$$

on peut démontrer que

$$\alpha_1 \wedge \cdots \wedge \alpha_N = \det[a_{ij}]\,dx_1 \wedge \cdots \wedge dx_N\,.$$

Prenons finalement $N = 3$ et

$$\alpha = A_1\,dx_1 + A_2\,dx_2 + A_3\,dx_3$$

$$\beta = B_1\,dx_2 \wedge dx_3 + B_2\,dx_3 \wedge dx_1 + B_3\,dx_1 \wedge dx_2\ ;$$

on vérifie que

$$\alpha \wedge \beta = A \cdot B\,dx_1 \wedge dx_2 \wedge dx_3$$

où $A = (A_1, A_2, A_3)$, $B = (B_1, B_2, B_3)$; ici

$$\alpha \wedge \beta = (-1)^{1\cdot 2}\beta \wedge \alpha = \beta \wedge \alpha$$

ce qui reflète la propriété de symétrie du produit scalaire $A \cdot B$.

Ainsi, on voit que le produit extérieur contient les règles de calcul des produits scalaires, des produits vectoriels (dans $\mathbb{R}^3$) et des déterminants.

L'utilisation de l'adjectif «extérieur» remonte à Grassmann dans ses écrits de 1844-1862 (*Ausdehnungslehre*) qui voulait distinguer ce produit des autres types de produits (tels que «produits intérieurs», «produits régressifs», etc.) ; c'est Elie Cartan (père de Henri Cartan) qui a introduit le calcul des formes différentielles systématiquement (d'abord formellement, comme dans ce chapitre, dans son article important «Sur certaines expressions différentielles et le problème de Pfaff» en 1899, cf. ses Œuvres complètes, Partie II) en s'inspirant, entre autres de Grassmann. La terminologie «produit extérieur», «différentiation extérieure», etc. était établie à partir de 1930 par les publications importantes de Elie Cartan.

7.1.3 Différentiation extérieure des formes

Une forme dans Ω, une partie ouverte non vide de $\mathbb{R}^N$, est dite de ***classe*** C^k si toutes ses composantes sont de classe C^k dans Ω.

Si $f \in C^1(\Omega)$, on notera

$$\mathrm{d}f = \mathrm{D}_1 f \, \mathrm{d}x_1 + \cdots + \mathrm{D}_N f \, \mathrm{d}x_N \, ;$$

on dit que la ***différentielle extérieure*** de la 0-forme f est la 1-forme $\mathrm{d}f$. Plus généralement, soit ω une p-forme de classe C^1 dans Ω, notée

$$\omega = \sum_I f_I \, \mathrm{d}x_I$$

où $I = (i_1, \ldots, i_p)$, $1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant N$ et $\mathrm{d}x_I = \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$; nous utiliserons ce genre d'abréviation pour alléger les formules ; alors on définit la ***différentielle extérieure*** $\mathrm{d}\omega$ de ω par la formule

$$\mathrm{d}\omega = \sum_I \mathrm{d}f_I \wedge \mathrm{d}x_I$$

où $\mathrm{d}f_I$ est défini ci-dessus.

Notons que si ω est une p-forme de classe C^k dans Ω, $\mathrm{d}\omega$ est une $(p+1)$-forme de classe C^{k-1} dans Ω.

Les propriétés principales de la différentiation extérieure sont résumées comme suit.

- *Si α, β sont deux p-formes de classe C^1 dans Ω, on a*

$$\mathrm{d}(\alpha + \beta) = \mathrm{d}\alpha + \mathrm{d}\beta \, .$$

- *Si $f \in C^1(\Omega)$ et ω est une p-forme de classe C^1 dans Ω alors*

$$\mathrm{d}(f\omega) = (\mathrm{d}f) \wedge \omega + f(\mathrm{d}\omega) \, . \tag{7.1}$$

Pour $p = 0$, (7.1) est la règle de dérivation usuelle ; si ω est donnée par la fonction g de classe C^1, (7.1) équivaut à la formule

$$\operatorname{grad}(fg) = (\operatorname{grad} f)g + f(\operatorname{grad} g)\,.$$

Pour $p = 1$ aussi, on peut vérifier (7.1) rapidement avec un petit calcul. La démonstration générale de cette formule ainsi que de celles qui suivent seront reprises dans la suite.

- Nous généralisons (7.1) en considérant deux formes α, β *de classe* C^1 *dans* Ω, α *étant une p-forme et* β *une q-forme* ; alors

$$\mathrm{d}(\alpha \wedge \beta) = (\mathrm{d}\alpha) \wedge \beta + (-1)^p \alpha \wedge (\mathrm{d}\beta)\,. \tag{7.2}$$

- Nous venons maintenant à la *propriété fondamentale* de la différentiation extérieure ; *si* ω *est une p-forme de classe* C^2 *dans* Ω *alors*

$$\mathrm{d}(\mathrm{d}\omega) = 0\,. \tag{7.3}$$

La preuve de (7.3) utilise, outre les règles de calcul énumérées précédemment, le fait que les dérivées mixtes d'une fonction de classe C^2 sont égales.

En prenant $N = 3$, nous allons vérifier que la différentiation extérieure redonne les opérations rot, div de l'analyse vectorielle. Soit

$$\omega = F_1\,\mathrm{d}x_1 + F_2\,\mathrm{d}x_2 + F_3\,\mathrm{d}x_3$$

une 1-forme de classe C^1 dans Ω ; on vérifie que

$$\begin{aligned}\mathrm{d}\omega = {} & (\mathrm{D}_2F_3 - \mathrm{D}_3F_2)\,\mathrm{d}x_2 \wedge \mathrm{d}x_3 + (\mathrm{D}_3F_1 - \mathrm{D}_1F_3)\,\mathrm{d}x_3 \wedge \mathrm{d}x_1 \\ & + (\mathrm{D}_1F_2 - \mathrm{D}_2F_1)\,\mathrm{d}x_1 \wedge \mathrm{d}x_2\end{aligned}$$

de sorte que les composantes de $\mathrm{d}\omega$ sont celles de $\operatorname{rot} F$ où $F = (F_1, F_2, F_3)$. Si l'on pose

$$\omega = F_1\,\mathrm{d}x_2 \wedge \mathrm{d}x_3 + F_2\,\mathrm{d}x_3 \wedge \mathrm{d}x_1 + F_3\,\mathrm{d}x_1 \wedge \mathrm{d}x_2$$

on aura

$$\mathrm{d}\omega = (\mathrm{D}_1F_1 + \mathrm{D}_2F_2 + \mathrm{D}_3F_3)\,\mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3\ ;$$

noter que la composante de $\mathrm{d}\omega$ est donnée par $\operatorname{div} F$ où $F = (F_1, F_2, F_3)$ est comme auparavant de classe C^1 dans Ω.

Pour $N = 2$ aussi on retrouve une expression familière du théorème de Green : si

$$\omega = F_1\,\mathrm{d}x_1 + F_2\,\mathrm{d}x_2$$

alors

$$\mathrm{d}\omega = (\mathrm{D}_1F_2 - \mathrm{D}_2F_1)\,\mathrm{d}x_1 \wedge \mathrm{d}x_2$$

si F_1, F_2 sont dans $C^1(\Omega)$.

En prenant $\omega = f \in C^2(\Omega)$, $\Omega \subset \mathbb{R}^3$, dans (7.3), on vérifie aisément que la relation (7.3) dans ce cas est équivalente à l'équation

$$\operatorname{rot}(\operatorname{grad} f) = 0\,.$$

De même, en prenant dans (7.3)

$$\omega = F_1\,\mathrm{d}x_1 + F_2\,\mathrm{d}x_2 + F_3\,\mathrm{d}x_3\,,$$

les F_i étant dans $C^2(\Omega)$, on vérifie que, dans ce cas, (7.3) équivaut à l'équation

$$\operatorname{div}(\operatorname{rot} F) = 0$$

où $F = (F_1, F_2, F_3)$. Ainsi, on retrouve deux identités importantes de l'analyse vectorielle comme cas particuliers de (7.3).

7.1.4 Formes fermées, formes exactes

De nouveau, Ω est une partie ouverte de $\mathbb{R}^N$. Une p-forme ω dans Ω s'appelle ***fermée*** dans Ω si ω est de classe C^1 dans Ω et $\mathrm{d}\omega = 0$ dans Ω ; elle s'appelle ***exacte*** s'il existe une $(p-1)$-forme π de classe C^1 dans Ω (p étant $\geqslant 1$) telle que $\mathrm{d}\pi = \omega$; dans ce cas, ω est une p-forme continue dans Ω (c'est-à-dire, chaque composante de ω est continue dans Ω) et on appelle π une ***primitive*** de ω.

De la formule (7.3) on voit que *toute forme exacte ω de classe C^1 dans Ω est fermée dans Ω*. En effet, la primitive π de ω est alors de classe C^2 dans Ω et

$$\mathrm{d}\omega = \mathrm{d}(\mathrm{d}\pi) = 0$$

selon (7.3).

Soit $F = (F_1, \ldots, F_N)$ un champ vectoriel de classe C^1 dans Ω ; associons à F la 1-forme ω dans Ω donnée par

$$\omega = F_1\,\mathrm{d}x_1 + \cdots + F_N\,\mathrm{d}x_N\,.$$

Un calcul montre que

$$\mathrm{d}\omega = \sum_{1 \leqslant i < j \leqslant N} (\mathrm{D}_i F_j - \mathrm{D}_j F_i)\,\mathrm{d}x_i \wedge \mathrm{d}x_j\,,$$

ainsi, *ω est une 1-forme fermée si et seulement si*

$$\mathrm{D}_i F_j = \mathrm{D}_j F_i = 0$$

dans Ω, pour $i \neq j$, $1 \leqslant i, j \leqslant N$, une condition déjà rencontrée dans le paragraphe 6.1.9 comme une condition *nécessaire* pour que F dérive d'un potentiel. D'autre part, *la 1-forme ω sera exacte, selon notre définition, si et seulement*

s'il existe une fonction u dans Ω (0-forme dans le présent langage) *de classe C^2 telle que*

$$\mathrm{d}u = \omega$$

dans Ω ; ceci équivaut à la relation

$$\operatorname{grad} u = F$$

dans Ω (c'est-à-dire u est un potentiel scalaire pour F dans le langage du paragraphe 6.1.9). On voit donc que la partie (ii) de la proposition 6.1.15 peut maintenant être énoncée comme suit : *si Ω est un domaine étoilé, une 1-forme ω de classe C^1 dans Ω est exacte si et seulement si ω est fermée.* Dans la suite, nous verrons que *cette proposition reste valable pour une p-forme ω* ; nous laissons comme exercice la vérification que la validité de cette proposition pour $p = 2$ dans un domaine étoilé $\Omega \subset \mathbb{R}^3$ est équivalente à la partie (ii) de la proposition 6.6.3 concernant l'existence d'un potentiel vecteur pour un champ incompressible dans Ω.

La terminologie « forme fermée » était introduite par de Rham dans son article de 1929 dont le contenu faisait partie ensuite de sa thèse de 1931 (cf. ses Œuvres mathématiques) ; les formes exactes s'appelaient à l'époque « formes homologues à zéro » ; l'origine de cette terminologie est dans la topologie combinatoire (appelée à l'époque « *analysis situs* »). Les travaux de de Rham ont montré les liens profonds entre la topologie algébrique et le calcul des formes différentielles via le théorème de Stokes généralisé.

7.1.5 Changement de variables dans les formes différentielles

Soient $U \subset \mathbb{R}^N$, $V \subset \mathbb{R}^M$ deux ensembles ouverts et $\varphi : U \to V$ une fonction de classe C^1 avec $\varphi = (\varphi_1, \ldots, \varphi_M)$. Notons par $x_1, \ldots, x_N$ les variables de $\mathbb{R}^N$ et par $y_1, \ldots, y_M$ celles de $\mathbb{R}^M$. Soit

$$\omega = \sum_I f_I \, \mathrm{d}y_I$$

une p-forme dans V (où $I = (i_1, \ldots, i_p)$, $1 \leqslant i_1 < \ldots < i_p \leqslant M$ et $\mathrm{d}y_I = \mathrm{d}y_{i_1} \wedge \cdots \wedge \mathrm{d}y_{i_p}$) ; on obtient une p-forme $\varphi^*\omega$ dans U en posant

$$\varphi^*\omega = \sum_I (f_I \circ \varphi) \, \mathrm{d}\varphi_I$$

où

$$\mathrm{d}\varphi_I = \mathrm{d}\varphi_{i_1} \wedge \cdots \wedge \mathrm{d}\varphi_{i_p} .$$

Ainsi $\varphi^*\omega$ s'obtient de ω en substituant φ_i en place de y_i, $i = 1, \ldots, M$; cette définition est possible même si φ est seulement dérivable ; si φ est de classe C^1, $\varphi^*\omega$ sera continue si ω est continue ; $\varphi^*\omega$ sera de classe C^{k-1} si ω et φ sont de classe C^k, $k \geqslant 1$. Nous dirons que $\varphi^*\omega$ est obtenue de ω par le changement

de variables $y = \varphi(x)$ où, plus brièvement, $\varphi^*\omega$ est l'***image réciproque*** de ω par φ ; la terminologie standard en anglais est que $\varphi^*\omega$ est le « *pullback* » de ω via φ.

Résumons ici les propriétés de φ^*.

- φ^* *commute avec toutes les opérations algébriques sur les formes dans* V :

$$\varphi^*(\alpha + \beta) = \varphi^*(\alpha) + \varphi^*(\beta)$$

si α, β sont deux p-formes dans V :

$$\varphi^*(a\alpha) = (a \circ \varphi)\,\varphi^*\alpha$$

si a est une 0-forme (une fonction réelle) dans V et α est une p-forme dans V ;

$$\varphi^*(\alpha \wedge \beta) = (\varphi^*\alpha) \wedge (\varphi^*\beta)$$

si α est une p-forme et β est une q-forme dans V.

- φ^* *commute avec* d :

$$\mathrm{d}(\varphi^*\omega) = \varphi^*(\mathrm{d}\omega)$$

si ω est une p-forme dans V de classe C^1 et φ *est de classe* C^2.

- *Transitivité des changements de variables* : si

$$U \xrightarrow{\varphi} V \xrightarrow{\psi} W$$

avec $U \subset \mathbb{R}^N$, $V \subset \mathbb{R}^M$ et $W \subset \mathbb{R}^P$, trois ensembles ouverts, φ, ψ étant de classe C^1 alors

$$(\psi \circ \varphi)^*\omega = \varphi^*\psi^*\omega$$

où ω est une p-forme dans W (de sorte que $(\psi \circ \varphi)^*\omega$ soit une p-forme dans U).

Les démonstrations de ces propriétés sont assez directes à partir de la définition des opérations φ^*, ψ^* et $(\psi \circ \varphi)^*$.

7.1.6 Intégration des formes différentielles : esquisse

Soit Ω une partie ouverte de $\mathbb{R}^N$. Une opération importante sur une p-forme ω dans Ω est celle de l'intégrer sur une partie d'une variété différentielle X de dimension p (de classe C^1) dans Ω ; si $p = 1$, il s'agit d'une intégrale curviligne ; si $p = 2$, il s'agit d'une intégrale de surface, la surface étant maintenant plongée dans $\mathbb{R}^N$ au lieu d'être dans $\mathbb{R}^3$ comme c'était le cas dans le chapitre précédent. Nous développerons la théorie générale dans la suite de ce chapitre ; indiquons ici seulement quelle est l'idée principale de ce developpement. Soit

M un morceau de la p-variété différentielle (contenu dans X mentionné ci-dessus) défini par un paramétrage $\psi : U \to X$, U étant un domaine dans $\mathbb{R}^p$ et $\psi(U) = M$; alors si

$$\omega = \sum_I f_I \, \mathrm{d}x_I$$

avec $\mathrm{d}x_I = \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$, $I = (i_1, \ldots, i_p)$, $1 \leqslant i_1 < \ldots < i_p \leqslant N$, on définit

$$\int_M \omega = \sum_I \int_U (f_I \circ \psi) \frac{\partial(\psi_{i_1}, \ldots, \psi_{i_p})}{\partial(u_1, \ldots, u_p)} \, \mathrm{d}u_1 \cdots \mathrm{d}u_p$$

où $\psi = (\psi_1, \ldots, \psi_N)$; en fait, cette définition explicite est équivalente à la suivante :

$$\int_M \omega = \int_U \psi^* \omega$$

où $\psi^*\omega$ est la p-forme dans U obtenue de ω dans Ω par le changement de variables $x = \psi(u)$; en effet,

$$\psi^* \omega = \sum_I (f_I \circ \psi) \, \mathrm{d}\psi_{i_1} \wedge \cdots \wedge \mathrm{d}\psi_{i_p}$$

et

$$\mathrm{d}\psi_{i_1} \wedge \cdots \wedge \mathrm{d}\psi_{i_p} = \frac{\partial(\psi_{i_1}, \ldots, \psi_{i_p})}{\partial(u_1, \ldots, u_p)} \, \mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_p$$

grâce à la formule pour $\alpha_1 \wedge \cdots \wedge \alpha_N$ dans le paragraphe 7.1.2. Puisque une p-forme γ dans U est de la forme

$$g \, \mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_p$$

où g est uniquement déterminée par γ, on peut définir

$$\int_U \gamma = \int_U g \, \mathrm{d}u$$

où l'intégrale à droite est l'intégrale de Lebesgue de g dans U. L'intégrale de ω sur un sous-domaine de la p-variété différentielle X est maintenant définie par un procédé de recollement que nous développerons dans la suite. Notre objectif est d'établir le ***théorème de Stokes*** généralisé dans la forme suivante : soit D un domaine régulier dans une p-variété différentielle (orientée) X contenue dans une partie ouverte Ω de $\mathbb{R}^N$; si ω est une $(p-1)$-forme de classe C^1 dans Ω alors

$$\int_{\partial D} \omega = \int_D \mathrm{d}\omega \, .$$

Il nous reste encore à définir : l'orientation de X, la notion du bord ∂D de D et l'orientation appropriée de ∂D. Le reste de ce chapitre est consacré à la précision des notions introduites ou indiquées dans cette section.

7.1.7 Exercices

Sauf mention du contraire, les formes sont définies dans une partie ouverte Ω de $\mathbb{R}^N$.

[1] Soient $\alpha = \sum_{i=1}^N f_i\,\mathrm{d}x_i$, $\beta = \sum_{i=1}^N g_i\,\mathrm{d}x_i$; vérifier que $\alpha \wedge \beta = 0$ si et seulement si $f_i g_j - f_j g_i = 0$ pour tout $i \neq j$; en particulier, $\alpha\wedge\alpha = 0$, $\beta\wedge\beta = 0$.

[2] Soient α une p-forme et β une q-forme ; vérifier que $\alpha\wedge\beta = \beta\wedge\alpha$ si p ou q est un nombre pair et $\alpha\wedge\beta = -(\beta\wedge\alpha)$ si p et q sont des nombres impairs. En déduire que $\alpha\wedge\alpha = 0$ si α est une p-forme avec p impair.

[3] Si $N = 3$, $\alpha\wedge\beta = 0$ pour deux 2-formes quelconques α et β ; plus généralement, si α est une p-forme et β est une q-forme et $p+q > N$, $\alpha\wedge\beta = 0$.

[4] Si $\omega = f\,\mathrm{d}x_1\wedge\mathrm{d}x_2 + g\,\mathrm{d}x_3\wedge\mathrm{d}x_4$ on a

$$\omega\wedge\omega = 2fg\,\mathrm{d}x_1\wedge\mathrm{d}x_2\wedge\mathrm{d}x_3\wedge\mathrm{d}x_4\,.$$

[5] Soit $N = 6$; posons $\omega = \mathrm{d}x_1\wedge\mathrm{d}x_4 + \mathrm{d}x_2\wedge\mathrm{d}x_5 + \mathrm{d}x_3\wedge\mathrm{d}x_6$. Vérifier que

$$\omega\wedge\omega\wedge\omega = -6\,\mathrm{d}x_1\wedge\mathrm{d}x_2\wedge\mathrm{d}x_3\wedge\mathrm{d}x_4\wedge\mathrm{d}x_5\wedge\mathrm{d}x_6\,.$$

[6] Soit $N = 2n$; posons $y_1 = x_{n+1}, \ldots, y_n = x_{2n}$ et

$$\omega = \mathrm{d}x_1\wedge\mathrm{d}y_1 + \cdots + \mathrm{d}x_n\wedge\mathrm{d}y_n\,.$$

Vérifier que

$$\begin{aligned}\omega^n &= (n!)\,\mathrm{d}x_1\wedge\mathrm{d}y_1\wedge\mathrm{d}x_2\wedge\mathrm{d}y_2\wedge\cdots\wedge\mathrm{d}x_n\wedge\mathrm{d}y_n\\ &= (-1)^{n(n-1)/2} n!\,\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_n\wedge\mathrm{d}y_1\wedge\cdots\wedge\mathrm{d}y_n\end{aligned}$$

où $\omega^n = \omega\wedge\cdots\wedge\omega$ (n fois). (Noter que $\mathrm{d}x_i\wedge\mathrm{d}y_i$ et $\mathrm{d}x_j\wedge\mathrm{d}y_j$ commutent.)

[7] Posons $\omega_i = \mathrm{d}x_1\wedge\cdots\widehat{\mathrm{d}x_i}\cdots\wedge\mathrm{d}x_N$ pour la $(N-1)$-forme où $\mathrm{d}x_i$ n'apparaît pas ; soit

$$\alpha = \sum_{i=1}^N (-1)^{i-1} f_i\omega_i$$

où les f_i sont de classe C^1. Vérifier que

$$\mathrm{d}\alpha = (\mathrm{D}_1 f_1 + \cdots + \mathrm{D}_N f_N)\,\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_N\,.$$

[8] Si α, β sont deux p-formes de classe C^1 alors $\mathrm{d}\alpha = \mathrm{d}\beta$ si et seulement si $(\alpha - \beta)$ est une forme fermée ; en particulier si $(\alpha - \beta)$ est exacte, $\mathrm{d}\alpha = \mathrm{d}\beta$.

9 Soit ω une 1-forme de classe C^1. Une fonction f, non nulle partout dans Ω, de classe C^1, s'appelle un *facteur intégrant* pour ω si $f\omega$ est fermé. Démontrer que l'existence d'un facteur intégrant pour ω entraîne que $\omega \wedge \mathrm{d}\omega = 0$. Pour $N = 3$, si $\omega = F_1\,\mathrm{d}x_1 + F_2\,\mathrm{d}x_2 + F_3\,\mathrm{d}x_3$, vérifier en détail que l'existence d'un facteur intégrant pour ω entraîne que $F \cdot \operatorname{rot} F = 0$ où $F = (F_1, F_2, F_3)$.

10 Soit $N = 2$ et $\Omega = \mathbb{R}^2 \setminus \{(0,0)\}$; posons

$$\widetilde{\Omega} = \{(r,\theta) : r > 0\,,\ -\pi < \theta < \pi\}\,, \quad \varphi(r,\theta) = (r\cos\theta\,,\, r\sin\theta)$$

$$\omega = \frac{-x_2}{x_1^2 + x_2^2}\,\mathrm{d}x_1 + \frac{x_1}{x_1^2 + x_2^2}\,\mathrm{d}x_2\,.$$

Pour $\varphi : \widetilde{\Omega} \to \Omega$, vérifier que $\varphi^*\omega = \mathrm{d}\theta$.

11 Soit U une partie ouverte de $\mathbb{R}^N$ et $\varphi : U \to \Omega$ une application surjective de classe C^1 ; notons les variables de U par $u_1, \cdots, u_N$. Si

$$\omega = f_1\,\mathrm{d}x_1 + \cdots + f_N\,\mathrm{d}x_N$$

est une 1-forme dans Ω, vérifier que

$$\varphi^*\omega = (\mathrm{D}_1\varphi \cdot f \circ \varphi)\,\mathrm{d}u_1 + \cdots + (\mathrm{D}_N\varphi \cdot f \circ \varphi)\,\mathrm{d}u_N$$

où $f = (f_1, \ldots, f_N)$. En conclure que si $\det \varphi'(u) \neq 0$, $u \in U$, alors $\varphi^*\omega = 0$ dans U si et seulement si $\omega = 0$ dans Ω.

12 (Continuation de l'exercice 11) Vérifier que $\varphi^*\omega = g\,\mathrm{d}u_N$ si et seulement si $\mathrm{D}_j\varphi \cdot (f \circ \varphi) = 0$, $j = 1, \ldots, N-1$.

13 (Continuation de l'exercice 11) Supposons que φ est de classe C^2 et $\det \varphi'(u) \neq 0$, $u \in U$; démontrer que ω est une 1-forme fermée (exacte) dans Ω si et seulement si $\varphi^*\omega$ est une 1-forme fermée (exacte) dans U.

14 Soit $\omega = \sum_I f_I\,\mathrm{d}x_I$ une p-forme avec $I = \{i_1 < \ldots < i_p\}$, $\mathrm{d}x_I = \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$, la somme étant sur tous les $I \subset \{1, \ldots, N\}$; posons $J = \{1, \ldots, N\} \setminus I = \{j_1 < \ldots < j_{N-p}\}$; soit π la permutation de $\{1, \ldots, N\}$ qui résulte en $\{i_1, \ldots, i_p, j_1, \ldots, j_{N-p}\}$ et $\varepsilon(\pi) = 1$ si π est paire, $\varepsilon(\pi) = -1$ si π est impaire. On définit la $(N-p)$-forme $*\omega$ par

$$*\omega = \sum_J \varepsilon(\pi) f_I\,\mathrm{d}x_J$$

($\omega \mapsto *\omega$ s'appelle l'*opérateur étoile* (de Hodge)).

(i) Vérifier que si $\omega = f_1\mathrm{d}x_1 + f_2\mathrm{d}x_2 + f_3\mathrm{d}x_3$ ($N = 3$), on a

$$*\,\omega = f_1\,\mathrm{d}x_2 \wedge \mathrm{d}x_3 + f_2\,\mathrm{d}x_3 \wedge \mathrm{d}x_1 + f_3\,\mathrm{d}x_1 \wedge \mathrm{d}x_2$$

$$**\omega = \omega\,.$$

(ii) Si $\omega = f_1\,dx_1 + f_2\,dx_2$ $(N = 2)$, on a

$$*\omega = f_1\,dx_2 - f_2\,dx_1\,, \quad **\omega = -\omega\,.$$

(iii) En général, pour une p-forme ω dans $\mathbb{R}^N$,

$$**\omega = (-1)^{p(N-p)}\omega\,.$$

7.2 PARTITIONS DE L'UNITÉ

7.2.1 Préliminaires

Généralisons d'abord la notion de fonctions à support compact introduite dans la section 5.11. Si $f : \Omega \to \mathbb{R}^M$ est une fonction quelconque définie dans une partie Ω de $\mathbb{R}^N$, nous définissons $\operatorname{supp} f$ comme l'adhérence de l'ensemble

$$A = \{x \in \Omega : f(x) \neq 0\}$$

dans Ω ; donc

$$\operatorname{supp} f = \overline{A} \cap \Omega\,;$$

nous dirons que f est à ***support compact*** si $\operatorname{supp} f$ est un ensemble compact. Ainsi, dire que $f : \Omega \to \mathbb{R}^M$ est une fonction à support compact équivaut à dire qu'il existe un ensemble compact K tel que $K \subset \Omega$ et $f(x) = 0$ si $x \notin K$.

Dans la section 5.11 on a vu l'existence de beaucoup de fonctions réelles de classe C^∞ dans $\mathbb{R}^N$ à support compact. Nous aurons besoin, en particulier, des fonctions suivantes : *si* $0 < a < b < \infty$ *et* $p \in \mathbb{R}^N$, *il existe une fonction* $f \in C^\infty(\mathbb{R}^N)$ *à support compact telle que* $0 \leqslant f \leqslant 1$ *et*

$$f(x) = 1 \text{ si } \|x - p\| \leqslant a\,, \quad f(x) = 0 \text{ si } \|x - p\| \geqslant b\,,$$
$$0 < f(x) < 1 \text{ si } a < \|x - p\| < b\,.$$

En effet, si $g \in C^\infty(\mathbb{R})$ telle que $0 \leqslant g \leqslant 1$, $g(t) = 0$ pour $t \leqslant a^2$, $g(t) = 1$ pour $t \geqslant b^2$ et $0 < g(t) < 1$ si $a^2 < t < b^2$, on obtient une fonction f du type désiré en posant

$$f(x) = 1 - g\Big(\|x - p\|^2\Big)\,, \quad x \in \mathbb{R}^N\,.$$

Pour avoir une fonction $g \in C^\infty(\mathbb{R})$ du type nécessaire, on utilise la fonction de Cauchy φ introduite dans la section 5.11 :

$$\varphi(t) = \mathrm{e}^{-1/t} \text{ si } t > 0\,, \quad \varphi(t) = 0 \text{ si } t \leqslant 0\,.$$

En posant

$$g(t) = \frac{\varphi(t-c)}{\varphi(t-c)+\varphi(d-t)}, \quad t \in \mathbb{R},$$

avec $-\infty < c < d < \infty$, on obtient $g \in C^\infty(\mathbb{R})$, $0 \leqslant g \leqslant 1$, et

$$g(t) = 0 \text{ si } t \leqslant c, \quad g(t) = 1 \text{ si } t \geqslant d, \quad 0 < g(t) < 1 \text{ si } c < t < d.$$

Notons que

$$\operatorname{supp} f = \overline{B}(p; b).$$

On écrit $f \in C_0^\infty(\mathbb{R}^N)$ si f est de classe C^∞ dans $\mathbb{R}^N$ et $\operatorname{supp} f$ est compact ; on dit que f est une fonction réelle de classe C^∞ dans $\mathbb{R}^N$ à support compact. De même, la notation $f \in C_0^\infty(\Omega)$ (où Ω est une partie ouverte de $\mathbb{R}^N$) signifie que f est une fonction réelle de classe C^∞ dans Ω à support compact dans Ω. Le sens des notations

$$f \in C_0^\infty(\mathbb{R}^N; \mathbb{R}^M), \quad f \in C_0^\infty(\Omega; \mathbb{R}^M)$$

doit être évident ; il s'agit d'une fonction à valeurs dans $\mathbb{R}^M$ de classe C^∞ dans $\mathbb{R}^N$, respectivement dans Ω, ayant un support compact.

7.2.2 Proposition

Soit $\{U_\alpha\}_{\alpha \in I}$ une famille de parties ouvertes de $\mathbb{R}^N$ avec $\Omega = \bigcup_{\alpha \in I} U_\alpha$. Alors, il existe une suite de fonction réelles $\{\theta_j\}$, $j = 1, 2, \ldots$, ayant les propriétés suivantes :

(i) on a $\theta_j \in C_0^\infty(\mathbb{R}^N)$, $0 \leqslant \theta_j \leqslant 1$, $j \geqslant 1$; pour chaque j, il existe un $\alpha_j \in I$ tel que $\operatorname{supp} \theta_j$ est contenu dans U_{α_j} et $\Omega = \bigcup_{j \geqslant 1} U_{\alpha_j}$;

(ii) si K est une partie compacte de $\mathbb{R}^N$ contenue dans Ω, il existe un entier $k \geqslant 1$ et une partie ouverte V dans Ω avec $K \subset V$ tels que, pour $n \geqslant k$,

$$\theta_1(x) + \cdots + \theta_n(x) = 1, \quad x \in V ;$$

(iii) pour tout $x \in \Omega$,

$$\sum_{j=1}^{\infty} \theta_j(x) = 1$$

et pour $x \in \Omega^c$, $\sum_j \theta_j(x) = 0$.

DÉMONSTRATION. Soit D un ensemble dénombrable contenu dans Ω et dense dans Ω ; ceci veut dire que $\overline{D} \supset \Omega$; on peut, par exemple, choisir D comme les points $x = (x_1, \ldots, x_N)$ de Ω tels que les x_i sont des nombres rationnels. Soit

$\{B_1, B_2, \ldots\}$ une énumération des boules fermées $\overline{B}(p;r)$ telles que $p \in D$, r est un nombre rationnel > 0 et

$$\overline{B}(p;r) \subset U_\alpha$$

pour un $\alpha \in I$; nous écrivons $B_n = \overline{B}(p_n;r_n)$, $n \geqslant 1$. Posons

$$V_n = B(p_n;\, r_n/2)\,, \quad n \geqslant 1\,;$$

vérifions que

$$\bigcup_{n=1}^{\infty} V_n = \Omega\,.$$

En effet, si $x \in \Omega$, il existe un $\alpha_x \in I$ tel que $x \in U_{\alpha_x}$ et ensuite il existe un $\varepsilon_x > 0$ tel que

$$B(x;\varepsilon_x) \subset U_{\alpha_x}$$

et un $p^x \in D$ tel que $\|x - p^x\| < (1/4)\,\varepsilon_x$; alors, si $(1/2)\,\varepsilon_x < t < (3/4)\,\varepsilon_x$, on a

$$x \in B(p^x;t/2) \subset B(p^x;t) \subset \overline{B}(p^x;t) \subset B(x;\varepsilon_x) \subset U_{\alpha_x}$$

car

$$\|x - p^x\| < \frac{1}{4}\,\varepsilon_x < \frac{t}{2}$$

et, si $\|y - p^x\| \leqslant t$, on a

$$\|y - x\| \leqslant \|y - p^x\| + \|p^x - x\| < t + \frac{1}{4}\,\varepsilon_x < \varepsilon_x\,.$$

Puisque $B(p^x;t/2)$ est un des V_n si t est un nombre rationnel entre $(1/2)\,\varepsilon_x$ et $(3/4)\,\varepsilon_x$, on a établi que Ω est une réunion des V_n.

Rappelons que $V_n = B(p_n;r_n/2)$; selon le paragraphe 7.2.1 il existe une fonction $f_n \in C_0^\infty(\mathbb{R}^N)$, $0 \leqslant f_n \leqslant 1$, avec

$$f_n(x) = 1 \text{ si } x \in \overline{V}_n\,, \quad f_n(x) = 0 \text{ si } \|x - p_n\| \geqslant r_n$$

$$0 < f_n(x) < 1 \quad \text{si } \frac{1}{2}\,r_n < \|x - p_n\| < r_n\,.$$

Posons

$$\theta_1 = f_1\,,\ \theta_2 = (1 - f_1)\,f_2\,,\ \theta_3 = (1 - f_1)(1 - f_2)\,f_3\,,\ \cdots$$

Il est clair que $\theta_j \in C_0^\infty(\mathbb{R}^N)$, $0 \leqslant \theta_j \leqslant 1$ et

$$\{x : \theta_j(x) \neq 0\} \subset \{x : f_j(x) \neq 0\} = B(p_j;r_j) \subset \overline{B}(p_j;r_j) \subset U_{\alpha_j}$$

pour un α_j approprié de sorte que $\operatorname{supp}\theta_j$ est contenu dans U_{α_j}.

Par récurrence, on vérifie que, pour $n = 1, 2, \ldots$,

$$\theta_1 + \cdots + \theta_n = 1 - \prod_{i=1}^{n}(1 - f_i)\,.$$

Si $x \in V_m$ on a $f_m(x) = 1$ de sorte que, pour $n \geqslant m$,

$$\sum_{j=1}^{n} \theta_j(x) = 1 \quad \text{si } x \in V_1 \cup \cdots \cup V_m\ ;$$

puisque Ω est la réunion des V_j, ceci établit (iii) de l'énoncé ; noter que $\theta_j(x) = 0$ pour tout j, si $x \in \Omega^c$. Si K est une partie compacte de $\mathbb{R}^N$ contenue dans Ω, il existe un k tel que

$$K \subset V_1 \cup \cdots \cup V_k = V\ ;$$

de ce qui précède, on voit que, pour $n \geqslant k$,

$$\theta_1(x) + \cdots + \theta_n(x) = 1\,, \quad x \in V\,. \qquad \blacklozenge$$

7.2.3 Remarques

Nous dirons que les $\{\theta_j\}_{j\geqslant 1}$ de la proposition 7.2.2 forment une ***partition de l'unité*** (dans Ω) ***subordonnée*** à la famille de parties ouvertes $\{U_\alpha\}_{\alpha\in I}$.

Les définitions d'une partition de l'unité utilisées par différents auteurs sont légèrement différentes l'une de l'autre ; mais dans toutes les définitions, les θ_j sont $\geqslant 0$, la somme de la famille $\{\theta_j\}$ est 1 (dans notre cas, la famille est une suite, éventuellement finie) et, dans un voisinage ouvert V approprié de chaque point x, la somme est, en fait, une somme d'un nombre fini k (dépendant de V) de θ_j (ceci résultant de la partie (ii) de la proposition 7.2.2). Notons que si Ω est une réunion finie des V_j de notre démonstration, alors on aura une suite finie de fonctions θ_j.

Une conséquence importante des remarques précédentes est que si $J \subset \{1, 2, \ldots\}$ et si

$$f(x) = \sum_{j\in J} \theta_j(x)\,, \quad x \in \Omega\,,$$

alors $f \in C^\infty(\Omega)$, $0 \leqslant f \leqslant 1$ et que

$$\mathrm{D}^\alpha f(x) = \sum_{j\in J} \mathrm{D}^\alpha \theta_j(x)\,, \quad x \in \Omega\,,$$

pour tout multi-indice α ; en effet, dans un voisinage ouvert approprié V de chaque point x la somme représentant f est, en fait, une somme d'un nombre fini de termes d'où la possibilité évidente de dériver terme à terme autant de fois que l'on veut.

Notons aussi que la somme des fonctions θ_j (de classe C^∞ dans $\mathbb{R}^N$) donne une fonction, en général, discontinue ; en effet, on a

$$\sum_{j=1}^{\infty} \theta_j = 1_\Omega$$

et 1_Ω est une fonction discontinue aux points $x \in \operatorname{Fr} \Omega$; rappelons que $\operatorname{Fr} \Omega \neq 0$ dès que Ω est une partie ouverte non vide différente de $\mathbb{R}^N$.

Nous illustrons maintenant l'utilisation de la proposition 7.2.2 dans la démonstration de différents énoncés utiles en analyse. Dans la suite, nous verrons que cette proposition nous sera précieuse pour la définition des intégrales dans une variété différentielle.

7.2.4 Proposition

Supposons que $F \subset U \subset \Omega$ où U et Ω sont deux parties ouvertes de $\mathbb{R}^N$ et F est une partie fermée dans Ω. Alors il existe une fonction $f \in C^\infty(\Omega)$, $0 \leqslant f \leqslant 1$, telle que

$$f(x) = \begin{cases} 1 & \text{si } x \in F \\ 0 & \text{si } x \in \Omega \setminus U. \end{cases}$$

DÉMONSTRATION. Soit $\{\theta_j\}_{j \geqslant 1}$ une partition de l'unité dans Ω, subordonnée à $\{U_1 = U\,,\ U_2 = \Omega \setminus F\}$. Posons

$$J_1 = \{j : \operatorname{supp} \theta_j \subset U_1\}\,, \quad J_2 = \{j : \operatorname{supp} \theta_j \subset U_2\}\,;$$

il est clair que $J_1 \cup J_2 = \{1, 2, \ldots\}$ et que

$$1 = \sum_{j \in J_1} \theta_j(x) + \sum_{j \in J_2} \theta_j(x) - \sum_{j \in J_1 \cap J_2} \theta_j(x)\,, \quad x \in \Omega\,,$$

Définissons

$$f(x) = \sum_{j \in J_1} \theta_j(x)\,, \quad x \in \Omega\,;$$

puisque $x \in F$ entraîne que $\theta_j(x) = 0$ pour $j \in J_2$ on voit que

$$f(x) = 1\,, \quad x \in F\,;$$

de même, $x \in \Omega \setminus U$ entraîne que $\theta_j(x) = 0$ pour $j \in J_1$ de sorte que $f(x) = 0$ pour $x \in \Omega \setminus U$; selon le paragraphe 7.2.3, $f \in C^\infty(\Omega)$ et $0 \leqslant f \leqslant 1$. ♦

7.2.5 Remarques

En prenant $\Omega = \mathbb{R}^N$ dans la proposition 7.2.4 on obtient $f \in C^\infty(\mathbb{R}^N)$, $0 \leqslant f \leqslant 1$, telle que

$$f(x) = 1 \text{ si } x \in F\,, \quad f(x) = 0 \text{ si } x \in U^c$$

où $F \subset U$, F fermé, U ouvert.

Par un raisonnement différent, on peut démontrer ce qui suit : *pour toute partie fermée non vide $F \subset \mathbb{R}^N$, il existe $f \in C^\infty(\mathbb{R}^N)$ telle que $F = f^{-1}(0)$.* Esquissons une démonstration ; soit $U = F^c$ et écrivons (en écartant le cas banal $U = \emptyset$)

$$U = \bigcup_{n=1}^{\infty} B(p_n; r_n)$$

pour un choix approprié des $p_n \in U$, $r_n > 0$; soit $f_n \in C_0^\infty(\mathbb{R}^N)$ telle que $0 \leqslant f_n \leqslant 1$ et telle que $f_n(x) > 0$ si et seulement si $x \in B(p_n; r_n)$; il suffit de prendre

$$f_n(x) = \varphi\Big(r_n^2 - \|x - p_n\|^2\Big)\,, \quad x \in \mathbb{R}^N\,,$$

avec φ fonction de Cauchy du paragraphe 7.2.1. Posons

$$f(x) = \sum_{n=1}^{\infty} \varepsilon_n f_n(x)\,, \quad x \in \mathbb{R}^N\,,$$

avec $\varepsilon_n > 0$ tel que

$$\varepsilon_n \cdot \sup\Big\{\big|\mathrm{D}^\alpha f_n(x)\big| : x \in \mathbb{R}^N\,,\ |\alpha| \leqslant n\Big\} \leqslant \frac{1}{2^n}\,, \quad n = 1, 2, \ldots\,;$$

un tel choix de ε_n est évidemment possible car $\sup_x \big|\mathrm{D}^\alpha f_n(x)\big| < \infty$ pour chaque α et il y a un nombre fini de multi-indices α tels que $|\alpha| \leqslant n$. La série définissant f est uniformément convergente dans $\mathbb{R}^N$ ainsi que les séries

$$\sum_{n=1}^{\infty} \varepsilon_n\, \mathrm{D}^\alpha f_n(x)$$

car les termes sont majorés par 2^{-n} pour $n \geqslant k$ si $|\alpha| \leqslant k$; donc (prop. 2.5.1) $f \in C^\infty(\mathbb{R}^N)$ et $f(x) = 0$ si $x \in F$ car $f_n(x) = 0$ pour tout n si $x \in F$; de plus, si $x \in F^c = U$, il existe un n tel que $f_n(x) > 0$ de sorte que $f(x) > 0$ si $x \in F^c$; aussi,

$$f(x) \leqslant \sum_{n=1}^{\infty} \varepsilon_n \leqslant \sum_{n=1}^{\infty} 2^{-n} = 1\,, \quad x \in \mathbb{R}^N\,. \qquad \blacklozenge$$

On trouvera d'autres résultats de ce type dans les exercices du paragraphe 7.2.9.

7.2.6 Proposition

Soit $A \subset \mathbb{R}^N$; supposons que $f : A \to \mathbb{R}^M$ est localement de classe C^k (voir la définition ci-dessous). Alors il existe une partie ouverte V contenant A et une fonction $\widetilde{f} \in C^k(V; \mathbb{R}^M)$ telle que $\widetilde{f}\big|_A = f$.

DÉMONSTRATION. La fonction $f : A \to \mathbb{R}^M$ est *localement de classe* C^k si pour tout $x \in A$, il existe un voisinage ouvert U_x de x et une fonction $f_x \in C^k(U_x; \mathbb{R}^M)$ tels que

$$f_x(y) = f(y) \quad \text{si } y \in A \cap U_x .$$

Posons $V = \bigcup_{x \in A} U_x$ et prenons une partition de l'unité $\{\theta_j\}_{j \geqslant 1}$ dans V, subordonnée à la famille $\{U_x\}_{x \in A}$; si $\operatorname{supp} \theta_j \subset U_{x_j}$, posons

$$\tilde{f}(x) = \sum_{j=1}^{\infty} f_j(x), \quad x \in V ,$$

où

$$f_j(x) = \begin{cases} \theta_j(x) f_{x_j}(x) & \text{si } x \in \operatorname{supp} \theta_j \\ 0 & \text{sinon.} \end{cases}$$

pour $j = 1, 2, \ldots$; noter que la somme donnée est, en fait, une somme finie dans un voisinage approprié de chaque $x \in V$; de plus, chaque f_j est évidemment de classe C^k dans $\mathbb{R}^N$ de sorte que $\tilde{f}$ est de classe C^k dans V. Rappelons aussi, que

$$A \subset V = \bigcup_{j=1}^{\infty} U_{x_j} ;$$

si $y \in A$, on vérifie que $f_j(y) = \theta_j(y) f(y)$ pour tout j de sorte que

$$\tilde{f}(y) = \sum_{j=1}^{\infty} \theta_j(y) \cdot f(y) = f(y) . \qquad \blacklozenge$$

7.2.7 Proposition

Soit $f : A \to \mathbb{R}^M$ une fonction continue définie dans une partie non vide quelconque A de $\mathbb{R}^N$. Si $\varepsilon : A \to]0, \infty[$ est telle que $\inf_{x \in K} \varepsilon(x) > 0$ pour toute partie compacte K contenue dans A alors il existe une fonction $\psi \in C^\infty(V; \mathbb{R}^M)$ définie dans un voisinage ouvert V de A telle que $|f(x) - \psi(x)| < \varepsilon(x)$, $x \in A$.

DÉMONSTRATION. Soit $x \in A$; puisque f est continue au point x, il existe une boule ouverte $B(x; t(x))$, telle que

$$|f(y) - f(x)| < \varepsilon(x), \quad y \in B(x; t(x)) \cap A ;$$

posons

$$\varepsilon_x = \inf \{\varepsilon(y) : y \in \overline{B}(x; t(x)) \cap A\} ;$$

on peut choisir $t(x)$ tel que $\varepsilon_x > 0$; il existe $r(x)$, $0 < r(x) \leqslant t(x)$, tel que

$$|f(y) - f(x)| < \varepsilon_x \leqslant \varepsilon(y), \quad y \in B(x; r(x)) = B_x , \; y \in A .$$

Posons $V = \bigcup_{x\in A} B_x$ et prenons une partition de l'unité $\{\theta_j\}_{j\geqslant 1}$ dans V, subordonnée à la famille $\{B_x\}_{x\in A}$; si $\operatorname{supp}\theta_j \subset B_{x_j}$, posons

$$\psi(x) = \sum_{j=1}^{\infty} \theta_j(x) f(x_j)\,, \quad x \in V\,.$$

Noter que la somme donnée est une somme finie dans un voisinage approprié de chaque $x \in V$; donc $\psi \in C^\infty(V;\mathbb{R}^M)$. De plus, pour $x \in V$,

$$\begin{aligned} \big|f(x) - \psi(x)\big| &= \left|\sum_{j=1}^{\infty} \theta_j(x)\big\{f(x) - f(x_j)\big\}\right| \\ &\leqslant \sum_{j=1}^{\infty} \theta_j(x)\big|f(x) - f(x_j)\big| < \varepsilon(x)\sum_{j=1}^{\infty}\theta_j(x) = \varepsilon(x) \end{aligned}$$

car

$$\theta_j(x)\big|f(x) - f(x_j)\big| \begin{cases} < \varepsilon(x)\theta_j(x) & \text{si } x \in \operatorname{supp}\theta_j \\ = \varepsilon(x)\theta_j(x) = 0 & \text{si } x \notin \operatorname{supp}\theta_j. \end{cases}$$

♦

La proposition suivante et son corollaire sont quelquefois très utiles.

7.2.8 Proposition

Soient U et V deux parties ouvertes non vides de $\mathbb{R}^N$ telles que $d(U,V) \geqslant \varepsilon > 0$. Il existe une fonction $f \in C^\infty(\mathbb{R}^N)$ telle que $0 \leqslant f \leqslant 1$, $f(x) = 1$ si $x \in U$, $f(x) = 0$ si $x \in V$ et

$$\big|\mathrm{D}_j f(x)\big| \leqslant \frac{M}{\varepsilon}\,, \quad x \in \mathbb{R}^N\,,\ j = 1, \ldots, N\,,$$

où M est une constante positive indépendante de U, V et ε.

DÉMONSTRATION. Posons $U_{\varepsilon/2} = \{x : d(x,U) < \varepsilon/2\}$; soit ψ une fonction dans $C_0^\infty(\mathbb{R}^N)$ telle que $\psi \geqslant 0$, $\psi(x) = 0$ si $\|x\| \geqslant 1$ et $\int_{\mathbb{R}^N}\psi = 1$ (§ 5.11.1) ; posons

$$\max_{1\leqslant j\leqslant N}\int_{\mathbb{R}^N}\big|\mathrm{D}_j\psi(x)\big|\,\mathrm{d}x = \frac{M}{2}\,.$$

Définissons f par la formule

$$f(x) = \left(\frac{\varepsilon}{2}\right)^{-N}\int_{U_{\varepsilon/2}}\psi\left(\frac{x-y}{\varepsilon/2}\right)\mathrm{d}y\,, \quad x \in \mathbb{R}^N\,.$$

On a

$$\begin{aligned} f(x) &= \left(\frac{\varepsilon}{2}\right)^{-N}\int_{\mathbb{R}^N} 1_{U_{\varepsilon/2}}(y)\psi\left(\frac{x-y}{\varepsilon/2}\right)\mathrm{d}y \\ &= \int_{\mathbb{R}^N} 1_{U_{\varepsilon/2}}\Big(x - \frac{\varepsilon}{2}z\Big)\psi(z)\,\mathrm{d}z\,, \quad x \in \mathbb{R}^N\,. \end{aligned}$$

Il résulte de ces formules que

$$f \in C^\infty(\mathbb{R}^N), \quad 0 \leqslant f \leqslant 1 ;$$

si $x \in U$, on aura $\big(x - (\varepsilon/2)z\big) \in U_{\varepsilon/2}$ dès que $\|z\| < 1$ d'où

$$f(x) = \int_{\{z:\|z\|<1\}} \psi(z)\,\mathrm{d}z = 1, \quad x \in U ;$$

si $x \in V$, on vérifie que

$$1_{U_{\varepsilon/2}}(y)\psi\left(\frac{x-y}{\varepsilon/2}\right) = 0, \quad y \in \mathbb{R}^N ;$$

en effet, ou bien $\|x-y\| \geqslant \varepsilon/2$ et alors $\psi\big((x-y)/(\varepsilon/2)\big) = 0$ ou bien $\|x-y\| < \varepsilon/2$ et alors $y \notin U_{\varepsilon/2}$ si $x \in V$, car $d(U,V) \geqslant \varepsilon$; ainsi, $f(x) = 0$ si $x \in V$. Finalement,

$$\mathrm{D}_j f(x) = \left(\frac{\varepsilon}{2}\right)^{-N} \int_{\mathbb{R}^N} 1_{U_{\varepsilon/2}}(y)\,\mathrm{D}_j\psi\left(\frac{x-y}{\varepsilon/2}\right)\frac{2}{\varepsilon}\,\mathrm{d}y, \quad x \in \mathbb{R}^N,$$

d'où

$$\big|\mathrm{D}_j f(x)\big| \leqslant \frac{2}{\varepsilon}\int_{\mathbb{R}^N}\big|\mathrm{D}_j\psi(z)\big|\,\mathrm{d}z \leqslant \frac{2}{\varepsilon}\cdot\frac{M}{2} = \frac{M}{\varepsilon}$$

pour $x \in \mathbb{R}^N$. ♦

COROLLAIRE. Soit A une partie compacte non vide de $\mathbb{R}^N$. Pour $\delta > 0$, il existe $f \in C^\infty(\mathbb{R}^N)$ tel que $0 \leqslant f \leqslant 1$, $f(x) = 0$ si $d(x,A) < \delta/2$, $f(x) = 1$ si $d(x,A) > \delta$ et $\big|\mathrm{D}_j f(x)\big| \leqslant 2M/\delta$, $x \in \mathbb{R}^N$, $1 \leqslant j \leqslant N$.

En effet, posons

$$U = \big\{x : d(x,A) > \delta\big\}, \quad V = \left\{x : d(x,A) < \frac{\delta}{2}\right\} ;$$

on a $U \subset A^c$, $V \supset A$ et $d(U,V) \geqslant \delta/2$ car si $u \in U$, $v \in V$,

$$\|u-v\| \geqslant \|u-a\| - \|a-v\| > \delta - \frac{\delta}{2} = \frac{\delta}{2}$$

en prenant $a \in A$ tel que $\|v-a\| < \delta/2$. La fonction f de la proposition avec $\varepsilon = \delta/2$ répond aux exigences du corollaire.

7.2.9 Remarques

La proposition 7.2.6 s'appelle un *théorème de prolongement* et la proposition 7.2.7 se dit un *théorème d'approximation*. Parmi d'autres théorèmes de ce type figurent le théorème de prolongement de Tietze-Urysohn vu dans le paragraphe 1.9.3 et le théorème d'approximation de Weierstrass vu dans la

proposition 5.11.5. Mentionnons spécialement les travaux de Whitney dans ce domaine (cf. ses *Collected Works* en deux volumes) ; ses résultats sur ce sujet (obtenus entre 1930 et 1960) ont eu des répercussions imprévues dans d'autres domaines comme celui de la théorie des catastrophes. Une introduction à la partie mathématique de cette dernière théorie peut être trouvée dans le livre de Bröcker et Lander [6] ; la proposition du paragraphe 7.2.5, due à Whitney, se trouve à la page 24 de cet ouvrage et notre démonstration de celle-ci suit leur présentation. Voir aussi Narasimhan [29, chap. 1] pour un exposé clair de plusieurs autres théorèmes de prolongement et d'approximation dus à Whitney.

7.2.10 Exercices

[1] Vérifier que si $f \in C_0^\infty(\mathbb{R}^N)$ et $g \in C^\infty(\mathbb{R}^N)$ alors $fg \in C_0^\infty(\mathbb{R}^N)$.

[2] Soient $f \in C^\infty(\mathbb{R}^N)$, $g \in C^\infty(U)$ où U est une partie ouverte de $\mathbb{R}^N$ et $f(x) = 0$ si $x \in U^c$. Vérifier que si g est bornée $fg \in C(\mathbb{R}^N)$ mais que fg n'est pas nécessairement dérivable. Qu'en est-il si g n'est pas bornée ?

[3] Soit $f \in C(\mathbb{R}^N)$ avec $\operatorname{supp} f = K$ compact ; si U est une partie ouverte de $\mathbb{R}^N$ telle que $K \subset U$ et $\varepsilon > 0$, démontrer qu'il existe une fonction $\psi \in C_0^\infty(\mathbb{R}^N)$ avec $\operatorname{supp} \psi \subset U$ telle que

$$\sup_x |f(x) - \psi(x)| < \varepsilon .$$

[4] Soient F_1, F_2 deux ensembles fermés disjoints non vides dans $\mathbb{R}^N$; si a, b sont deux nombres réels distincts quelconques, démontrer qu'il existe une fonction bornée $f \in C^\infty(\mathbb{R}^N)$ telle que

$$f^{-1}(a) = F_1 , \quad f^{-1}(b) = F_2 .$$

(Utiliser la proposition du paragraphe 7.2.5 et une construction donnée dans le paragraphe 1.9.1.)

[5] Soit $f : \mathbb{R}^N \to \mathbb{R}$ une fonction continue et $\varepsilon > 0$; démontrer qu'il existe une fonction $\psi \in C^\infty(\mathbb{R}^N)$ telle que

$$\int_{\mathbb{R}^N} |f(x) - \psi(x)| < \varepsilon .$$

(Utiliser la proposition 7.2.7 avec $\varepsilon(x)$ approprié.)

[6] Soit $f \in C(\mathbb{R}^2)$ une fonction telle que $f(x) \to 0$ lorsque $\|x\| \to \infty$. Démontrer que, pour tout $\varepsilon > 0$, il existe un n et des fonction $g_i \in C_0^\infty(\mathbb{R})$, $h_i \in C_0^\infty(\mathbb{R})$, $i = 1, \ldots, n$, tels que

$$\sup_{(x_1,x_2)\in\mathbb{R}^2} \left| f(x_1, x_2) - \sum_{i=1}^{n} g_i(x_1) h_i(x_2) \right| < \varepsilon .$$

Généraliser le résultat à $\mathbb{R}^N$.

7 Démontrer qu'il existe $f : \mathbb{R} \to \mathbb{R}^N$ de classe C^∞ telle que $f(\mathbb{R})$ est dense dans $\mathbb{R}^N$.
(Prendre $\{p_1, p_2, \ldots\}$ dense dans $\mathbb{R}^N$; définir $g \in C(\mathbb{R}, \mathbb{R}^N)$ telle que $g(n) = p_n$, $n = 1, 2, \ldots$; selon la proposition 7.2.7, il existe $f : \mathbb{R} \to \mathbb{R}^N$ de classe C^∞ telle que $|g(t) - f(t)| < 1/(1+t^2)$, $t \in \mathbb{R}$; si $x \in \mathbb{R}^N$, $\varepsilon > 0$, trouver p_n avec $n \geqslant n_0$ tel que $|x - p_n| < \varepsilon/2$ où n_0 est tel que $n^{-2} < \varepsilon/2$ pour $n \geqslant n_0$; alors $|x - f(n)| < \varepsilon$.)

8 Démontrer que si $f : \mathbb{R} \to \mathbb{R}^N$, $n > 1$, est de classe C^1 alors $f(\mathbb{R})$ ne peut pas contenir une partie ouverte non vide. (Noter que $\lambda_N(f(\mathbb{R})) = 0$; cf. § 5.6.6).

7.3 INTÉGRATION DANS LES VARIÉTÉS DIFFÉRENTIELLES

7.3.1 Préliminaires

Rappelons d'abord la notion d'une sous-variété différentielle de $\mathbb{R}^N$, de dimension k et de classe C^r, r étant un entier $\geqslant 1$ ou $r = \infty$ et $k = 0, 1, \ldots, N$; cette notion a été introduite dans le chapitre 4, spécialement dans le paragraphe 4.4.2 et la proposition 4.4.3. Nous reprenons la discussion du chapitre 4 dans une notation légèrement modifiée pour la rapprocher de notre présentation des surfaces du chapitre précédent.

Une partie non vide $E \subset \mathbb{R}^N$ s'appelle un ***morceau d'une variété différentielle*** (ou simplement ***morceau***) de dimension $k \geqslant 1$ et de classe C^r s'il existe un domaine $U \subset \mathbb{R}^k$ et une application $\psi : U \to \mathbb{R}^N$ tels que :

(i) $\psi(U) = E$ et ψ est un homéomorphisme entre U et E ;
(ii) ψ est de classe C^r ;
(iii) pour tout $u \in U$, $\operatorname{rang} \psi'(u) = k$.

Dans la terminologie du paragraphe 4.4.2, ψ est un plongement (de classe C^r) de U dans $\mathbb{R}^N$; notons que notre partie ouverte U dans $\mathbb{R}^k$ est un domaine (c'est-à-dire U est non vide, ouvert et connexe) ; ceci n'est pas essentiel mais il correspond à notre définition précédente d'un morceau de surface (qui est le cas $k = 2$) dans $\mathbb{R}^3$. Nous dirons que (ψ, U) (ou plus brièvement ψ) est un ***paramétrage*** du morceau E, la variable u dans U étant le ***paramètre***.

Une partie non vide X de $\mathbb{R}^N$ s'appelle une ***variété différentielle*** (de dimension $k \geqslant 1$, de classe C^r) si tout point $x \in X$ possède un voisinage ouvert V dans $\mathbb{R}^N$ tel que $V \cap X$ est un morceau de dimension k de classe C^r. Si (ψ, U) est un paramétrage du morceau $V \cap X$ on dit que (ψ, U) est un ***paramétrage local*** de X dans un voisinage de x ; l'ensemble $V \cap X = V_x$ est un voisinage X-ouvert du point $x \in X$; l'application réciproque $\psi^{-1} : V_x \to U$ s'appelle une ***carte locale*** du voisinage V_x de x (étant sous-entendu ici que V_x est X-ouvert

et que $x \in V_x$) ; si $\psi^{-1} = \varphi = (\varphi_1, \ldots, \varphi_k)$, $\big(\varphi_1(y), \ldots, \varphi_k(y)\big)$, $y \in V_x$, s'appelle un système de ***coordonnées locales*** défini dans V_x. Une famille de ***paramétrages locaux*** $(\psi_\alpha, U_\alpha)_{\alpha \in I}$ s'appelle un ***atlas*** pour X si

$$\bigcup_{\alpha \in I} \psi_\alpha(U_\alpha) = X\,.$$

Si $\psi_\alpha(U_\alpha) = V_\alpha$ et $\varphi_\alpha = \psi_\alpha^{-1}$, certains auteurs appellent $(\varphi_\alpha, V_\alpha)_{\alpha \in I}$ un atlas pour X ; mais nous n'adopterons pas cette terminologie.

Rappelons qu'une variété différentielle X dans $\mathbb{R}^N$ s'appelle aussi une sous-variété différentielle de $\mathbb{R}^N$, la justification de cette terminologie étant que $\mathbb{R}^N$ est lui-même une variété différentielle de dimension N dans $\mathbb{R}^N$ (avec un paramétrage $\psi : \mathbb{R}^N \to \mathbb{R}^N$, $\psi(x) \equiv x$). La terminologie permet aussi l'usage suivant : « soit X une sous-variété différentielle de la variété différentielle Y dans $\mathbb{R}^N$ » ; ceci veut dire que $X \subset Y$ et X, Y tous les deux sont des variétés différentielles (de dimensions et classes non spécifiées).

Rappelons aussi qu'une sous-variété différentielle de dimension 0 dans $\mathbb{R}^N$ (la classe C^r ne sera pas mentionnée dans ce contexte) est une partie discrète $X \subset \mathbb{R}^N$ dans le sens que tout point $x \in X$ possède un voisinage ouvert V dans $\mathbb{R}^N$ tel que $V \cap X = \{x\}$.

Même si nous travaillons généralement avec les paramétrages locaux, il est profitable de se rappeler la caractérisation des variétés différentielles données dans la proposition 4.4.3 en utilisant les cartes redressantes. Une partie non vide X de $\mathbb{R}^N$ est une sous-variété différentielle de dimension $k \geqslant 1$ et de classe C^r si et seulement si pour tout point $a \in X$ on a une carte redressante (f, W) où W est un voisinage ouvert de a dans $\mathbb{R}^N$, f est un C^r-difféomorphisme entre W et $f(W)$, un voisinage ouvert de $0 \in \mathbb{R}^N$, avec $f(a) = 0$ et

$$f(W \cap X) = \big\{y = (y_1, \ldots, y_N) \in f(W) : y_{k+1} = \cdots = y_N = 0\big\}\ ;$$

ceci a comme conséquence que la partie $W \cap X$ est décrite par les équations de la forme

$$f_{k+1}(x) = 0\,,\ \ldots\ ,\ f_N(x) = 0$$

où $f = (f_1, \ldots, f_N)$. La démonstration que l'on a donnée consiste à associer un paramétrage local à chaque carte redressante et vice versa. En effet, étant donnée une carte redressante (f, W) au point $a \in X$, posons

$$U = \big\{u = (u_1, \ldots, u_k) : (u_1, \ldots, u_k, 0, \ldots, 0) \in f(W)\big\}\ ;$$

pour $u \in U$, il existe un unique $x \in W \cap X$ tel que $f(x) = u$; en définissant $\psi(u) = x$, on obtient un paramétrage local (ψ, U) de X (de classe C^r) dans le voisinage $W \cap X$ de a ; nous dirons que ce paramétrage local ψ est ***associé*** à la carte redressante (f, W) au point $a \in X$. Le paramétrage local associé ψ est tel que $\psi^{-1} = f\big|_{W \cap X}$ est aussi de classe C^r de sorte que ψ est un C^r-difféomorphisme entre $U \subset \mathbb{R}^k$ et $\psi(U) = W \cap X \subset \mathbb{R}^N$. En fait, comme

pour les surfaces, *tout paramétrage* (ψ, U) *(de classe* C^r*) de* X *dans un voisinage* V_a *de* $a \in X$ *est un* C^r*-difféomorphisme*. En effet, la démonstration de la proposition 4.4.3 montre qu'étant donné (ψ, U), la fonction $\psi^{-1} : V_a \to U$ (qui est déjà un homéomorphisme) est de classe C^r localement autour de chaque point $x \in V_a$ d'où il résulte que ψ^{-1} est de classe C^r dans V_a (cf. prop. 7.2.6).

Comme pour les morceaux de surfaces, chaque morceau de dimension $k \geqslant 1$ (de classe C^r) dans $\mathbb{R}^N$ possède localement une représentation sous la forme du graphe d'une fonction à valeurs dans $\mathbb{R}^{N-k}$ de k variables réelles. Il est clair que

$$(x_1, \ldots, x_k) \longmapsto (x_1, \ldots, x_k, \varphi_1(x_1, \ldots, x_k), \ldots, \varphi_{N-k}(x_1, \ldots, x_k))$$

avec $(x_1, \ldots, x_k) \in U$ donne un morceau (de classe C^r) de dimension k si $\varphi = (\varphi_1, \ldots, \varphi_{N-k})$ est de classe C^r dans U ; inversement, si (ψ, U) est un paramétrage local de X dans un voisinage de a (de classe C^r) on peut résoudre les équations

$$\psi_1(u_1, \ldots, u_k) = x_1\,, \ \ldots\,, \ \psi_N(u_1, \ldots, u_k) = x_N$$

localement pour obtenir

$$x_{k+1} = \varphi_1(x_1, \ldots, x_k)\,, \ \ldots\,, \ x_N = \varphi_{N-k}(x_1, \ldots, x_k)$$

pourvu que

$$\frac{\partial(\psi_1, \ldots, \psi_k)}{\partial(u_1, \ldots, u_k)} \neq 0\,,$$

ce que l'on peut supposer, sans perte de généralité, car $\operatorname{rang} \psi' = k$. Nous n'élaborons pas ce raisonnement car il est parfaitement analogue à celui donné dans le cas des surfaces dans le chapitre 6.

Une variété différentielle X de dimension $k \geqslant 1$ dans $\mathbb{R}^N$ est dite de ***codimension*** $N-k$; si la codimension est 1, on dit que la variété différentielle est une ***hypersurface*** dans $\mathbb{R}^N$.

Dans la suite, nous ne mentionnerons plus explicitement la classe C^r d'une variété différentielle ; r sera fixé une fois pour toute et il sera mentionné seulement si sa valeur joue un rôle particulier. Par conséquent, on parlera seulement de paramétrage local, sans mentionner sa classe, car il sera tacitement supposé être de la classe C^r fixée. Aussi, la dimension k d'une variété différentielle sera supposée être $\geqslant 1$ si rien d'autre n'est précisé.

7.3.2 Changement de paramétrage local

Soit X une variété différentielle de dimension k dans $\mathbb{R}^N$. Si (ψ_1, U_1), (ψ_2, U_2) sont deux paramétrages locaux de deux voisinages V_1, V_2 d'un même point $a \in X$, posons

$$V = V_1 \cap V_2\,, \quad W_1 = \psi_1^{-1}(V)\,, \quad W_2 = \psi_2^{-1}(V)\,.$$

Alors, $\varphi = \psi_2^{-1} \circ \psi_1$ *est un* C^r*-difféomorphisme de* W_1 *sur* W_2 car ψ_2^{-1} et ψ_1 sont de classe C^r (cf. la discussion du § 7.3.1). Nous supposons que V_1, V_2 et V sont connexes de sorte que W_1, W_2 sont connexes aussi ; donc la quantité $\det \varphi'(u)$, qui est toujours non nulle dans W_1, ne change pas de signe. Si $\det \varphi'(u) > 0$ pour tout $u \in W_1$, on dira que le changement de paramétrage de ψ_1 à ψ_2 donné par φ (ou vice versa) est ***direct*** ; si $\det \varphi'(u) < 0$ pour tout $u \in W_1$, on dira que le changement de paramétrage de ψ_1 à ψ_2 (ou vice versa) donné par φ est ***opposé***.

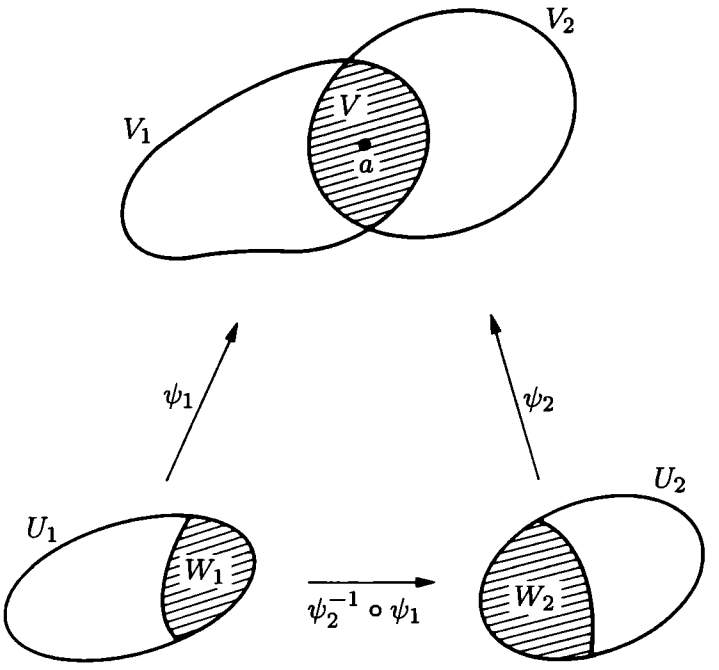


Fig. 7.1

La discussion précédente s'applique, en particulier, au cas où l'on a deux paramétrage (ψ_1, U), (ψ_2, U) basés sur un même domaine $U \subset \mathbb{R}^k$; alors $\varphi = \psi_2^{-1} \circ \psi_1$ sera un C^r-difféomorphisme d'une partie de U sur lui-même (qui peut être direct ou opposé).

Le fait que $\varphi = \psi_2^{-1} \circ \psi_1$ est un difféomorphisme est décrit souvent dans les termes suivants : les coordonnées locales se transforment différentiablement lorsque l'on passe d'une carte locale à une autre.

7.3.3 Espaces tangents

Soit X une variété différentielle de dimension k dans $\mathbb{R}^N$. Soit $a \in X$; si (ψ, U) est un paramétrage local d'un voisinage V_1 de $a \in X$, avec $\psi(\alpha) = a$, nous définissons

$$T_a(X) = T_a = \psi'(\alpha)(\mathbb{R}^k) \subset \mathbb{R}^N \ ;$$

T_a s'appelle l'***espace tangent*** à X en a ; c'est un sous-espace vectoriel de $\mathbb{R}^N$ de dimension k car $\operatorname{rang} \psi'(a) = k$. *L'espace* T_a *est indépendant du paramétrage*

local (ψ, U) *choisi pour sa définition.* En effet, si (ψ_2, U_2) est un autre paramétrage local d'un autre voisinage V_2 de a avec $\psi_2(\beta) = a$, on a vu (§ 7.2.2) que $\varphi = \psi_2^{-1} \circ \psi$ est un difféomorphisme de W_1 sur W_2 où

$$a \in V = V_1 \cap V_2\,, \quad \alpha \in W_1 = \psi^{-1}(V_1)\,, \quad W_2 = \psi_2^{-1}(V_2)\,;$$

puisque $\varphi(\alpha) = \beta$ et $\psi_2 \circ \varphi = \psi$ on obtient

$$\psi'(\alpha) = \psi_2'(\beta)\,\varphi'(\alpha)\,;$$

or, $\varphi'(\alpha)$ étant un isomorphisme de $\mathbb{R}^k$ sur $\mathbb{R}^k$, on a que

$$\psi'(\alpha)(\mathbb{R}^k) = \psi_2'(\beta)(\mathbb{R}^k)\,;$$

ce qui établit notre affirmation.

Comme dans le cas des surfaces dans $\mathbb{R}^3$, on peut démontrer sans peine l'affirmation suivante : *un vecteur* $\xi \in \mathbb{R}^N$, $\xi \neq 0$, *appartient à* T_a *si et seulement s'il existe un chemin régulier* $\gamma : [-1, 1] \to X$, $\gamma(0) = a$, $\gamma'(0) = \xi$.

L'espace $T_a^\perp$ dans $\mathbb{R}^N$ s'appelle l'***espace normal*** à X en a ; on a

$$\dim T_a^\perp = N - k = \text{codimension de } X\,;$$

si $\xi \in T_a^\perp$, on dit que ξ est un ***vecteur normal*** (***unitaire***) à X en a (si $\|\xi\| = 1$) ; on dira aussi que ξ est une ***normale unitaire*** à X en a si $\xi \in T_a^\perp$, $\|\xi\| = 1$. Si X est une hypersurface (c'est-à-dire $k = N - 1$), il y a exactement deux normales unitaires à X en a données par $\pm\xi$ où $\xi \in T_a^\perp$, $\|\xi\| = 1$; si X est donnée par une équation

$$x_N = f(x_1, \ldots, x_{N-1})\,, \quad (x_1, \ldots, x_{N-1}) \in U$$

où U est un domaine dans $\mathbb{R}^{N-1}$ et $a = (a_1, \ldots, a_N) \in X$ alors T_a est le sous-espace de $\mathbb{R}^N$ engendré par les vecteurs

$$\xi_1 = \big(1, 0, \ldots, 0, \mathrm{D}_1 f(\alpha)\big)\,, \ldots,, \ \xi_{N-1} = \big(0, 0, \ldots, 1, \mathrm{D}_{N-1} f(\alpha)\big)$$

où $\alpha = (a_1, \ldots, a_{N-1})$. En effet, le paramétrage associé est

$$\psi(u_1, \ldots, u_{N-1}) = \big(u_1, \ldots, u_{N-1}, f(u_1, \ldots, u_{N-1})\big)$$

avec $(u_1, \ldots, u_{N-1}) \in U$; en écrivant $a = (\alpha; a_N)$ on voit que $\psi'(\alpha)$ est donné par la matrice du type $(N\,,\, N-1)$ suivante :

$$\psi'(\alpha) = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & 1 \\ \mathrm{D}_1 f(\alpha) & \mathrm{D}_2 f(\alpha) & & \mathrm{D}_{N-1} f(\alpha) \end{bmatrix} = \big[\xi_1, \ldots, \xi_{N-1}\big]$$

d'où notre affirmation concernant $T_a = \psi'(\alpha)(\mathbb{R}^{N-1})$. On en conclut que $T_a^\perp = \{c\xi : c \in \mathbb{R}\}$ où

$$\xi = \big(-\mathrm{D}_1 f(\alpha), \ldots, -\mathrm{D}_{N-1} f(\alpha), 1\big)\,;$$

en effet, si $\xi = (c_1, \ldots, c_N) \in T_a^{\perp}$ et $\xi \neq 0$ alors

$$c_1 + c_N \mathrm{D}_1 f(\alpha) = 0\,, \ \ldots\,, \ c_{N-1} + c_N \mathrm{D}_{N-1} f(\alpha) = 0$$

d'où $c_N = c \neq 0$ et $c_i = -c\,\mathrm{D}_i f(\alpha)$, $1 \leqslant i \leqslant N-1$, ce qui établit notre formule pour $T_a^{\perp}$. La normale unitaire

$$\nu(a) = \nu = \frac{\xi}{\|\xi\|}\,, \quad \|\xi\| = \left\{1 + \mathrm{D}_1^2 f(\alpha) + \cdots + \mathrm{D}_{N-1}^2 f(\alpha)\right\}^{1/2}$$

sera appelée dans la suite, la normale unitaire ***standard*** associée à X au point $a = (\alpha; a_N)$. Les mêmes considérations seront applicables si l'hypersurface X est donnée par une équation de type $x_i = f(x_1, \ldots, \widehat{x_i}, \ldots, x_N) = f(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N)$ (le chapeau $\widehat{\ }$ indiquant, en général, l'omission du symbole concerné).

7.3.4 Volume d'un parallélépipède

Soient $\xi_1, \ldots, \xi_k$, k vecteurs dans $\mathbb{R}^N$ où $k \leqslant N$; posons

$$A = [\xi_1, \ldots, \xi_k]$$

la matrice du type (N, k) formée par les k vecteurs colonnes $\xi_1, \ldots, \xi_k$; notons que

$${}^tAA = [\xi_i \cdot \xi_j]\,, \quad 1 \leqslant i, j \leqslant k\,.$$

Posons

$$P = P[\xi_1, \ldots, \xi_k]$$

le parallélépipède engendré par $\xi_1, \ldots, \xi_k$ dans $\mathbb{R}^N$; P est formé des vecteurs

$$t_1\xi_1 + \cdots + t_k\xi_k$$

où les $t_i \geqslant 0$, $\sum_{i=1}^k t_i = 1$. Dans le paragraphe 5.6.4, on a proposé dans (5.25) de définir la mesure k-dimensionnelle de P par la formule

$$\sigma_k(P) = \left\{\mathrm{gram}(\xi_1, \ldots, \xi_k)\right\}^{1/2} = \left\{\det({}^tAA)\right\}^{1/2}$$

où $\mathrm{gram}(\xi_1, \ldots, \xi_k)$ est le déterminant de Gram (ou le grammien) défini par les vecteurs $\xi_1, \ldots, \xi_k$:

$$\mathrm{gram}(\xi_1, \ldots, \xi_k) = \det[\xi_i \cdot \xi_k]\,.$$

Dans ce paragraphe, nous justifions cette définition par un argument *ad hoc*. Une discussion complète demande une définition de $\sigma_k(E)$ pour les ensembles boréliens $E \subset \mathbb{R}^N$ puis une vérification du fait que $\sigma_k(P)$ est effectivement donné par la formule proposée. Cette démarche est faisable de plusieurs manières, mais chacune d'elles est longue (cf. Federer [18, pp. 171-174]). Ici, nous nous contenterons de proposer un raisonnement convaincant.

Vérifions d'abord que

$$\operatorname{gram}(\xi_1, \ldots, \xi_k) \geqslant 0\,;$$

il suffit pour cela de considérer la forme quadratique engendrée par $\alpha_{ij} = \xi_i \cdot \xi_j$; on a que

$$\sum_{1 \leqslant i,\, j \leqslant k} \alpha_{ij} t_i t_j = \left(\sum_{i=1}^{k} t_i \xi_i\right) \cdot \left(\sum_{i=1}^{k} t_i \xi_i\right) \geqslant 0$$

pour tout choix de $t_i, \ldots, t_k$ réels de sorte que la matrice $[\alpha_{ij}]$ est de type positif d'où

$$\det[\alpha_{ij}] \geqslant 0$$

avec positivité stricte si et seulement si les vecteurs $\xi_1, \ldots, \xi_k$ sont linéairement indépendants. La positivité de $\det[\alpha_{ij}]$ résulte aussi du raisonnement qui suit.

Notons maintenant que si $k = N$, on a déjà vu dans le paragraphe 5.6.4, (5.24), que

$$\lambda_N(P) = \{\operatorname{gram}(\xi_1, \ldots, \xi_N)\}^{1/2} = \{\det({}^tAA)\}^{1/2} = |\det A|\,.$$

Si $k < N$ mais les vecteurs $\xi_1, \ldots, \xi_k$ sont dans $\mathbb{R}^k \times \{0\}$, de sorte que,

$$\xi_i = (\eta_i; 0)\,, \quad \eta_i \in \mathbb{R}^k,\ 0 \in \mathbb{R}^{N-k}\,,$$

pour $i = 1, \ldots, k$, on aura

$$A = [\xi_1, \ldots, \xi_k] = \begin{bmatrix} \eta_1, \ldots, \eta_k \\ 0, \ldots, 0 \end{bmatrix}$$

et

$${}^tAA = [\eta_i \cdot \eta_j]\,, \quad 1 \leqslant i\,,\ j \leqslant k\,.$$

Ainsi, $\{\det({}^tAA)\}^{1/2}$ peut être interprété comme le volume k-dimensionnel du parallélépipède engendré par $\eta_1, \ldots, \eta_k$ dans $\mathbb{R}^k$, en vue du résultat mentionné précédemment. Lorsque $\xi_1, \ldots, \xi_k$ $(k < N)$ sont quelconques, on peut trouver une transformation orthogonale $V : \mathbb{R}^N \to \mathbb{R}^N$ qui transforme le sous-espace engendré par $\xi_1, \ldots, \xi_k$ en un sous-espace de $\mathbb{R}^k \times \{0\}$; posons

$$V\xi_i = \widetilde{\xi}_i\,, \quad i = 1, \ldots, k\,,$$

et $\widetilde{A}$ la matrice formée par les colonnes $\widetilde{\xi}_1, \ldots, \widetilde{\xi}_k$. Alors, $\widetilde{A} = VA$ (V étant la matrice orthogonale du type (N, N) correspondante) et

$${}^t\widetilde{A}\widetilde{A} = {}^tA\,{}^tVVA = {}^tAA$$

de sorte que

$$\left\{\det({}^t\widetilde{A}\widetilde{A})\right\}^{1/2} = \{\det({}^tAA)\}^{1/2}$$

d'où, en admettant qu'une transformation orthogonale laisse les mesures k-dimensionnelles invariantes, on conclut que

$$\{\det({}^tAA)\}^{1/2} = \{\mathrm{gram}(\xi_1, \ldots, \xi_k)\}^{1/2}$$

donne bien la mesure k-dimensionnelle du parallélépipède engendré par $\xi_1, \ldots, \xi_k$ dans $\mathbb{R}^N$.

Notons aussi que pour $k = 1$ on obtient la longueur de ξ_1 et que pour $k = 2$, $N = 3$, on a la formule usuelle de l'aire d'un parallélogramme comme expliqué dans le paragraphe 6.3.3.

Avec $A = [\xi_1, \ldots, \xi_k]$ comme ci-dessus, il sera commode pour la suite d'avoir une formule alternative pour

$$\det({}^tAA) \ ;$$

on a la formule suivante :

$$\det({}^tAA) = \sum_I \det{}^2 A_I \tag{7.4}$$

où I parcourt tous les sous-ensembles $\{i_1 < \ldots < i_k\}$ dans $\{1, \ldots, N\}$ et A_I est la matrice de type (k, k) obtenue de A en retenant seulement ses lignes numéros $i_1, \ldots, i_k$. Une interprétation géométrique de (7.4) est la suivante : le carré du volume k-dimensionnel de P est égal à la somme des carrés des volumes des parallélépipèdes obtenus en projetant P sur les différents sous-espaces coordonnés de dimension k dans $\mathbb{R}^k$. La formule (7.4) sera déduite d'une formule plus générale appelée ***formule de Binet-Cauchy***. Notons un cas particulier de (7.4) que l'on a déjà vu. Avec $k = 2$, $N = 3$, on a

$$A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix} \ ;$$

en écrivant $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3)$, on obtient de (7.4) que

$$\det \begin{bmatrix} \|a\|^2 & a \cdot b \\ a \cdot b & \|b\|^2 \end{bmatrix} = (a_1b_2 - a_2b_1)^2 + (a_2b_3 - a_3b_2)^2 + (a_1b_3 - a_3b_1)^2$$

ce qui équivaut à

$$\|a\|^2 \|b\|^2 - (a \cdot b)^2 = \|a \times b\|^2 ,$$

c'est-à-dire à

$$\|a\|^2 \|b\|^2 = \|a \times b\|^2 + (a \cdot b)^2 .$$

La vérification « usuelle » de cette identité se fait en écrivant

$$a \cdot b = \|a\| \cdot \|b\| \cos\theta , \quad \|a \times b\| = \|a\| \cdot \|b\| \sin\theta .$$

7.3.5 Fomule de Binet-Cauchy

Soit $X = [x_{ij}]$ une matrice (m,n), $Y = [y_{ij}]$ une matrice (n,m) et $Z = XY = [z_{ij}]$ la matrice carrée du type (m,m) obtenue comme produit de X et Y. Alors $\det Z = 0$ si $m > n$; si $m \leqslant n$, on a

$$\det Z = \sum_{1 \leqslant j_1 < \ldots < j_m \leqslant n} X\begin{pmatrix} 1 & 2 & \ldots & m \\ j_1 & j_2 & \ldots & j_m \end{pmatrix} Y\begin{pmatrix} j_1 & j_2 & \ldots & j_m \\ 1 & 2 & \ldots & m \end{pmatrix} \tag{7.5}$$

où

$$X\begin{pmatrix} 1 & 2 & \ldots & m \\ j_1 & j_2 & \ldots & j_m \end{pmatrix}$$

est le déterminant de la matrice (m,m) obtenue de X en retenant seulement ses colonnes numéros $j_1, \ldots, j_m$ et

$$Y\begin{pmatrix} j_1 & j_2 & \ldots & j_m \\ 1 & 2 & \ldots & m \end{pmatrix}$$

est le déterminant de la matrice (m,m) obtenue de Y en retenant seulement ses lignes numéros $j_1, \ldots, j_m$.

Pour démontrer (7.5) on fait les calculs suivants : puisque

$$z_{ij} = \sum_{\alpha=1}^{n} x_{i\alpha} y_{\alpha j} ,$$

$$\begin{aligned} \det Z &= \det \begin{bmatrix} \sum_{\alpha_1=1}^{n} x_{1\alpha_1} y_{\alpha_1 1} & \cdots & \sum_{\alpha_m=1}^{n} x_{1\alpha_m} y_{\alpha_m m} \\ \vdots & & \vdots \\ \sum_{\alpha_1=1}^{n} x_{m\alpha_1} y_{\alpha_1 1} & & \sum_{\alpha_m=1}^{n} x_{m\alpha_m} y_{\alpha_m m} \end{bmatrix} \\ &= \sum_{1 \leqslant \alpha_1, \ldots, \alpha_m \leqslant n} \det \begin{bmatrix} x_{1\alpha_1} y_{\alpha_1 1} & \cdots & x_{1\alpha_m} y_{\alpha_m m} \\ \vdots & & \vdots \\ x_{m\alpha_1} y_{\alpha_1 1} & \cdots & x_{m\alpha_m} y_{\alpha_m m} \end{bmatrix} \\ &= \sum_{1 \leqslant \alpha_1, \ldots, \alpha_m \leqslant n} X\begin{pmatrix} 1 & 2 & & m \\ \alpha_1 & \alpha_2 & & \alpha_m \end{pmatrix} y_{\alpha_1 1}\, y_{\alpha_2 2} \cdots y_{\alpha_m m} . \end{aligned} \tag{7.6}$$

Si $m > n$, chaque terme du second membre de (7.6) est nul car au moins deux des α_i sont égaux ; donc, dans ce cas $\det Z = 0$ et (7.5) est établie.

Soit $m \leqslant n$; dans ce cas, on retient seulement les termes du second membre de (7.6) qui correspondent aux $\alpha_1, \ldots, \alpha_m$ distincts car les autres sont nuls ; on a donc que le second membre de (7.6) est égal à :

$$\sum_{1\leqslant\alpha_1<\ldots<\alpha_m\leqslant n}\sum_{\pi} X\begin{pmatrix}1 & \ldots & m\\ \pi(\alpha_1) & \ldots & \pi(\alpha_m)\end{pmatrix} y_{\pi(\alpha_1)1}\cdots y_{\pi(\alpha_m)m}$$

$$= \sum_{1\leqslant\alpha_1<\ldots<\alpha_m\leqslant n} X\begin{pmatrix}1 & \ldots & m\\ \alpha_1 & \ldots & \alpha_m\end{pmatrix}\sum_{\pi}\sigma(\pi) y_{\pi(\alpha_1)1}\cdots y_{\pi(\alpha_m)m}$$

$$= \sum_{1\leqslant\alpha_1<\ldots<\alpha_m\leqslant n} X\begin{pmatrix}1 & \ldots & m\\ \alpha_1 & \ldots & \alpha_m\end{pmatrix} Y\begin{pmatrix}\alpha_1 & \ldots & \alpha_m\\ 1 & \ldots & m\end{pmatrix}$$

ce qui donne (7.5) ; ici, π parcourt les $m!$ permutations de $\alpha_1 < \ldots < \alpha_m$ et $\varepsilon(\pi) = \pm 1$ est la signature de la permutation π (égale à $(-1)^\nu$, ν étant égal au nombre d'inversions d'ordre dans $\pi(\alpha_1), \ldots, \pi(\alpha_m)$).

Si $X = {}^tA$, $Y = A$, A étant une matrice (N, k), on a

$$X\begin{pmatrix}1 & \ldots & m\\ j_1 & \ldots & j_m\end{pmatrix} = {}^tA\begin{pmatrix}1 & \ldots & m\\ j_1 & \ldots & j_m\end{pmatrix} = A\begin{pmatrix}j_1 & \ldots & j_m\\ 1 & & m\end{pmatrix}$$

$$Y\begin{pmatrix}j_1 & \ldots & j_m\\ 1 & \ldots & m\end{pmatrix} = A\begin{pmatrix}j_1 & \ldots & j_m\\ 1 & \ldots & m\end{pmatrix}$$

et (7.5) donne (7.4) immédiatement.

Notons un cas particulier de (7.5) que nous avons déjà utilisé plusieurs fois : si $m = n$,

$$\det(XY) = \det(X)\det(Y)\,.$$

7.3.6 L'élément de volume

Soit E un morceau de dimension k dans $\mathbb{R}^N$ donné par un paramétrage (ψ, U). Posons, pour $u \in U$,

$$\rho_\psi(u) = \left[\det\left\{{}^t\psi'(u)\,\psi'(u)\right\}\right]^{1/2}$$

$$= \left[\sum_I \left\{\frac{\partial(\psi_{i_1}, \ldots, \psi_{i_k})}{\partial(u_1, \ldots, u_k)}\right\}^2 (u)\right]^{1/2} \tag{7.7}$$

où $u = (u_1, \ldots, u_k)$, $\psi = (\psi_1, \ldots, \psi_k)$, $I = \{i_1 < \ldots < i_k\}$ avec I parcourant toutes les suites du type indiqué de longueur k avec $1 \leqslant i_1 < \ldots < i_k \leqslant N$. La seconde égalité de (7.7) vient de la formule (7.4) du paragraphe 7.3.4. Nous écrivons ρ au lieu de ρ_ψ s'il n'y a pas d'ambiguïté. On définit une mesure positive $\mu = \sigma_k^E$ pour les parties boréliennes de E en posant

$$\int_E f\,\mathrm{d}\mu = \int_U (f \circ \psi)\,\rho_\psi\,\mathrm{d}u \tag{7.8}$$

où $f : E \to [0, \infty]$ est une fonction borélienne ; alors $\mu(A) = \int_E 1_A\,\mathrm{d}\mu$, $A \subset E$. Démontrons que *cette définition est indépendante du choix de paramétrage*

(ψ, U). En effet, si (γ, U) est un second paramétrage, on sait (§ 7.3.2) que $\varphi = \gamma^{-1} \circ \psi$ est un C^1-difféomorphisme de U sur U ; le calcul suivant :

$$\gamma \circ \varphi = \psi\,, \quad (\gamma' \circ \varphi)\varphi' = \psi'$$
$${}^t\psi'\psi' = {}^t\varphi'\{{}^t\gamma' \circ \varphi \cdot \gamma' \circ \varphi\}\varphi'$$

combiné avec la propriété multiplicative des déterminants donne que

$$\rho_\psi = \rho_\gamma \circ \varphi \cdot |\det \varphi'|$$

$$\int_U f \circ \psi \cdot \rho_\psi \, \mathrm{d}u = \int_U f \circ \gamma \circ \varphi \cdot \rho_\gamma \circ \varphi \cdot |\det \varphi'| \, \mathrm{d}u$$
$$= \int_U f \circ \gamma \cdot \rho_\gamma \, \mathrm{d}v$$

en posant $v = \varphi(u)$. Ceci établit que la définition dans la formule (7.8) est indépendante du choix de paramétrage.

La formule (7.8) donne immédiatement que

$$\sum_{n=1}^{\infty} \int_E f_n \, \mathrm{d}\mu = \int_E \sum_{n=1}^{\infty} f_n \, \mathrm{d}\mu$$

d'où, en posant

$$\mu(A) = \int_E 1_A \, \mathrm{d}\mu$$

pour $A \subset E$ (A borélien), on obtient la mesure positive μ qui sera notée, si nécessaire, par σ_k^E, pour indiquer qu'elle correspond à un volume k-dimensionnel des parties de E. On écrit aussi

$$\mathrm{d}\mu = \rho_\psi \, \mathrm{d}u$$

pour résumer (7.8) et on dit que $\rho = \rho_\psi$ définit l'***élément de volume*** dans E.

Si $k = 1$, $\mathrm{d}\mu$ définit l'intégrale par rapport à la longueur (notée alors $\mathrm{d}s$) ; si $k = 2$, $N = 3$, $\mathrm{d}\mu$ définit l'intégrale superficielle (notée alors $\mathrm{d}\sigma$) que l'on a déjà vu dans le chapitre 6.

La justification intuitive pour la formule

$$\mathrm{d}\mu = \rho_\psi \, \mathrm{d}u$$

est la suivante : on a vu (§ 7.3.4) que le parallélépipède engendré par les k vecteurs $\mathrm{D}_1\psi(u), \ldots, \mathrm{D}_k\psi(u)$ dans $\mathbb{R}^N$ possède le volume k-dimensionnel $\rho_\psi(u)$; un volume infinitésimal $\mathrm{d}u_1 \, \mathrm{d}u_2 \cdots \mathrm{d}u_k = \mathrm{d}u$ dans U est transformé par ψ dans un volume infinitésimal

$$\mathrm{d}\mu = \rho_\psi(u) \, \mathrm{d}u$$

dans E.

Lorsque le morceau E dans $\mathbb{R}^N$ est de dimension $N-1$ (cas d'un morceau d'une hypersurface) et E est présenté par une équation du type

$$x_N = f(x_1, \ldots, x_{N-1}), \quad (x_1, \ldots, x_{N-1}) \in U,$$

l'élément de volume dans E s'écrit comme

$$\mathrm{d}\mu = \{1 + \mathrm{D}_1^2 f + \cdots + \mathrm{D}_{N-1}^2 f\}^{1/2} \, \mathrm{d}x_1 \cdots \mathrm{d}x_{N-1}. \tag{7.9}$$

En effet, ici

$$\psi(u) = \big(u; f(u)\big), \quad u \in U,$$

où $u = (x_1, \ldots, x_{N-1})$ et

$$\psi'(u) = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{D}_1 f(u) & \mathrm{D}_2 f(u) & \ldots & \mathrm{D}_{N-1} f(u) \end{bmatrix}$$

une matrice de type $(N, N-1)$; de (7.4) avec $k = N-1$, on voit que

$$\det\big({}^t\psi'(u)\,\psi'(u)\big) = 1 + \mathrm{D}_1^2 f(u) + \cdots + \mathrm{D}_{N-1}^2 f(u)$$

d'où l'on a (7.9). Notons que l'analogue exact de (7.9) dans $\mathbb{R}^3$ a déjà été vu dans le chapitre 6 (§ 6.3.3). Noter que si $k = N$, un morceau de dimension k dans $\mathbb{R}^N$ est un domaine E de $\mathbb{R}^N$ et la formule (7.8) est simplement la formule du changement de variable dans les intégrales dans $\mathbb{R}^N$, avec $\mu = \lambda_N$; dans ce cas, l'élément de volume de E sera donné par $\mathrm{d}\mu = \mathrm{d}x$.

7.3.7 Mesure naturelle dans une variété

Soit X une sous-variété différentielle de dimension k $(1 \leqslant k \leqslant N)$ dans $\mathbb{R}^N$; dans ce paragraphe on définit une mesure positive $\mu = \sigma_k^X$ (ou simplement σ^X ou σ_k) pour les parties boréliennes de X par

$$\int_X f \, \mathrm{d}\mu$$

pour $f : X \to [0, \infty]$, une fonction borélienne positive. Si E est une partie de X qui est un morceau de dimension k et $A \subset E$, on aura

$$\mu(A) = \sigma^E(A)$$

où σ^E est défini comme dans le paragraphe 7.3.6 ; la définition complète de μ sera faite en utilisant une partition de l'unité appropriée.

Pour $x \in X$, soit E_x un morceau de dimension k tel que $x \in E_x \subset X$ avec $E_x = W_x \cap X$ où W_x est un voisinage ouvert de x dans $\mathbb{R}^N$. Soit $\{\theta_j\}_{j\geqslant 1}$ une partition de l'unité subordonnée à la famille $\{W_x\}_{x\in X}$; posons

$$\int_X f \,\mathrm{d}\mu = \sum_{j=1}^{\infty} \int f\theta_j \,\mathrm{d}\sigma^j \tag{7.10}$$

où $\operatorname{supp}\theta_j \subset W_{x_j}$, $E_j = E_{x_j} = W_{x_j} \cap X$ et σ^j est la mesure positive associée au morceau E_j, c'est-à-dire

$$\sigma^j = \sigma^{E_j} .$$

Noter que $f\theta_j = 0$ en dehors de E_j ; démontrons que *(7.10) est indépendante du choix de la partition de l'unité utilisée*. En effet, si $\{\widetilde{\theta}_j\}$, $W_{\tilde{x}_j}$ $\widetilde{E}_j = E_{\tilde{x}_j}$ $\widetilde{\sigma}^j = \sigma^{\widetilde{E}_j}$ correspondent à un autre choix, on aura

$$\begin{aligned}
\sum_{j=1}^{\infty} \int f\widetilde{\theta}_j \,\mathrm{d}\widetilde{\sigma}^j &= \sum_{j=1}^{\infty} \int \sum_{k=1}^{\infty} f\widetilde{\theta}_j\theta_k \,\mathrm{d}\widetilde{\sigma}^j \\
&= \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \int f\widetilde{\theta}_j\theta_k \,\mathrm{d}\sigma^k = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \int f\widetilde{\theta}_j\theta_k \,\mathrm{d}\sigma^k \\
&= \sum_{k=1}^{\infty} \int \sum_{j=1}^{\infty} f\widetilde{\theta}_j\theta_k \,\mathrm{d}\sigma^k = \sum_{k=1}^{\infty} \int f\theta_k \,\mathrm{d}\sigma^k
\end{aligned}$$

où l'on a utilisé le fait que $f\widetilde{\theta}_j\theta_k$ est zéro en dehors de $\widetilde{E}_j \cap E_k$ et que $\widetilde{\sigma}^j = \sigma^k$ pour les parties contenues dans $\widetilde{E}_j \cap E_k$; ceci établit l'indépendance de la définition donnée par (7.10) du choix des θ_j. Noter que toutes les opérations d'inversion des sommes et intégrales sont justifiées par la positivité des termes ; ceci établit aussi immédiatement que

$$\int_X \sum_{n=1}^{\infty} f_n \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_X f_n \,\mathrm{d}\mu$$

pour les $f_n : X \to [0, \infty]$ boréliennes. En posant,

$$\mu(A) = \int_X 1_A \,\mathrm{d}\mu$$

pour $A \subset X$, A borélien, on obtient une mesure positive définie pour les parties boréliennes de X. Pour $k = N$, cette mesure μ est la mesure de Lebesgue λ_N dans la partie ouverte X de $\mathbb{R}^N$.

Il est utile de définir la mesure positive 0-dimensionnelle pour une sous-variété de dimension zéro X comme la mesure de comptage ; rappelons qu'une

telle sous-variété est une partie discrète, au plus dénombrable, de $\mathbb{R}^N$; dans ce cas

$$\int_X f \, d\mu = \sum_{n=1}^{\infty} f(x_n)$$

si $X = \{x_1, x_2, \ldots\}$ et $f : X \to [0, \infty]$.

La mesure $\mu = \sigma^X = \sigma_k^X$ s'appellera la ***mesure naturelle*** dans la variété X de dimension k.

La mesure naturelle σ^k d'une sous-variété différentielle X de dimension 1 s'appelle aussi la ***longueur*** ; on vérifie facilement qu'une intégrale curviligne d'une fonction scalaire f,

$$\int_\gamma f \, ds$$

est la même chose que

$$\int_C f \, d\sigma^X$$

si γ est un chemin ou une trajectoire simple régulier dont la trace C est contenue dans la variété différentielle X de dimension 1 (en supposant évidemment que ces intégrales existent).

De même, la mesure naturelle σ^X d'une sous-variété différentielle de dimension 2 s'appelle l'***aire*** ; si $A \subset X$, $f : A \to \mathbb{R}$, on a l'écriture

$$\int_A f \, d\sigma = \int_A f \, d\sigma^X$$

en supposant toujours que ces intégrales existent.

Notons que si X est une sous-variété différentielle de dimension k et σ^X est sa mesure naturelle alors pour toute fonction $f : A \to \mathbb{R}$ (A partie borélienne de X et f une fonction réelle borélienne),

$$\int_A f \, d\sigma^X$$

existe dès que $\sigma^X(A) < \infty$ et $M = \sup_{x \in A} |f(x)| < \infty$ car

$$\int_A |f| \, d\sigma^X \leqslant M\sigma^X(A) < \infty ;$$

ce sera le cas si A est une partie compacte de X et f est continue dans A (Ex. 8, § 7.3.10). Evidemment, il y a bien d'autres cas aussi où f sera σ^X-intégrable dans $A \subset X$.

Il faut noter que $\sigma^X(X)$ peut être égal à ∞ même si X est une sous-variété différentielle bornée de dimension $1 \leqslant k < N$ dans $\mathbb{R}^N$ (Ex. 9, § 7.3.10) ; évidemment si X est compact alors $\sigma^X(X) < \infty$ (Ex. 8, § 7.3.10). Puisque une sous-variété différentielle X de dimension N dans $\mathbb{R}^N$ est une partie ouverte

de $\mathbb{R}^N$ avec $\sigma^X = \lambda_N$, on a dans ce cas, qu'une partie (borélienne) A de X (X étant de dimension N) est de mesure naturelle finie dès que A est bornée.

Le cas des courbes compactes C^1 par morceaux n'est pas directement traitée dans la discussion précédente ; mais une telle courbe est toujours de longueur finie, même si elle n'est pas contenue dans une variété différentielle de dimension 1, à cause de ses points multiples et de ses points irréguliers.

7.3.8 Ensembles négligeables dans une variété

Soit X une sous-variété différentielle de dimension k dans $\mathbb{R}^N$, $1 \leqslant k \leqslant N$. Une partie (borélienne) A de X s'appelle ***négligeable*** dans X si

$$\mu(A) = 0$$

où $\mu = \sigma^X$ est la mesure naturelle dans X.

Notons d'abord que $A \subset X$ *est négligeable dans* X *si et seulement si*

$$\mu(A \cap E) = 0$$

pour tout morceau E *de dimension* k *contenu dans* X. Ceci est évident car X est une réunion dénombrable des morceaux de dimension k contenus dans X.

Soit $f : D \to X$, D une partie ouverte de $\mathbb{R}^m$ avec $m < k$; *si* f *est de classe* C^1 *alors* $f(D)$ *est une partie négligeable dans* X. En effet, il suffit de voir que $f(D) \cap E$ est négligeable dans X pour un morceau E de dimension k contenu dans X ; supposons que E est paramétré par (ψ, U), U un domaine de $\mathbb{R}^k$. De la définition

$$\mu(B) = \int_U (1_B \circ \psi)\rho_\psi \, du = \int_{\psi^{-1}(B)} \rho_\psi \, du$$

pour $B \subset E = \psi(U)$, on voit que B est négligeable dans E si et seulement si $\psi^{-1}(B)$ est négligeable dans U car $\rho_\psi(u) > 0$ pour $u \in U$. Puisque E est une partie relativement ouverte dans X, l'ensemble $D_1 = f^{-1}(E)$ est une partie ouverte contenue dans D et

$$f(D_1) = f(D) \cap E \, ;$$

si $g = f|_{D_1}$, $\psi^{-1} \circ g$ est une application de classe C^1 de $D_1 \subset \mathbb{R}^m$ dans $U \subset \mathbb{R}^k$, $k > m$; du paragraphe 5.6.6 nous savons que $\lambda_k(\widetilde{N}) = 0$ où

$$\widetilde{N} = (\psi^{-1} \circ g)(D_1) \subset U \, ;$$

ceci entraîne que

$$\psi(\widetilde{N}) = g(D_1) = f(D) \cap E$$

est une partie négligeable dans E et conclut la preuve de notre affirmation.

Une conclusion immédiate de ce qui précède est la suivante : *si* $Y \subset X$ *et* Y *est une sous-variété différentielle de dimension* $m < k$ *alors* Y *est une partie négligeable dans* X.

7.3.9 Ensembles k-négligeables

Soit $A \subset \mathbb{R}^N$; pour $\varepsilon > 0$, posons

$$A_\varepsilon = \{x : d(x, A) \leqslant \varepsilon\} ;$$

notons que A_ε est un ensemble fermé ; si k est un entier, $1 \leqslant k \leqslant N$, nous dirons que A est ***k-négligeable*** si

$$\lim_{\varepsilon \to 0} \varepsilon^{-N+k} \lambda_N(A_\varepsilon) = 0 .$$

Les affirmations suivantes concernant les ensembles k-négligeables nous seront utiles.

(i) Un ensemble borné A *est* N*-négligeable si et seulement si* $\lambda_N(\overline{A}) = 0$.

En effet, les relations

$$\overline{A} = \bigcap_{n=1}^{\infty} A_{1/n} , \quad \lambda_N(\overline{A}) = \lim_{n \to \infty} \lambda_N(A_{1/n})$$

prouvent (i).

(ii) *Si* A *est* k*-négligeable et* $B \subset A$ *alors* B *est* k*-négligeable.*

(iii) *Si* $A_1, \ldots, A_n$ *sont* k*-négligeables alors* $A_1 \cup \cdots \cup A_n$ *est* k*-négligeable.*

Les preuves de (ii) et (iii) sont immédiates.

(iv) *Soit* ℓ *un entier tel que* $1 \leqslant \ell < k \leqslant N$ *; si* A *est une partie compacte de* $\mathbb{R}^\ell$ *et* $\varphi : A \to \mathbb{R}^N$ *est une application lipschitzienne alors* $\varphi(A)$ *est* k*-négligeable.*

DÉMONSTRATION. Supposons que, pour $\alpha > 0$,

$$|\varphi(u) - \varphi(v)| \leqslant \alpha |u - v| , \quad u , v \in A ;$$

Soit $T > 0$ tel que $A \subset [-T , T]^\ell$; soit ε tel que $0 < \varepsilon < 2\alpha T$ et $\varepsilon' = \varepsilon/\alpha$; prenons un entier $n \geqslant 2$ tel que

$$\frac{\varepsilon'}{2} \leqslant \frac{2T}{n} < \varepsilon' .$$

Il existe $r \leqslant n^\ell$ points $u_1, \ldots, u_r$ dans A tels que pour tout $a \in A$,

$$\min_i |a - u_i| < \varepsilon' ;$$

pour le voir, on peut subdiviser $[-T , T]$ en n intervalles égaux donnant n^ℓ cubes dont la réunion est $[-T , T]^\ell$; il suffit alors de choisir un point de A de chacun des cubes qui intersectent A. Vérifions maintenant que

$$\left\{x \in \mathbb{R}^N : d(x, \varphi(A)) \leqslant \varepsilon\right\} \subset \bigcup_{i=1}^{r} \left\{x \in \mathbb{R}^N : |x - \varphi(u_i)| < 2\varepsilon\right\} ;$$

en effet, si $d(x,\varphi(A)) \leqslant \varepsilon$, on a un $a \in A$ tel que

$$\|x - \varphi(a)\| \leqslant \varepsilon$$

d'où $|x - \varphi(a)| \leqslant \|x - \varphi(a)\| \leqslant \varepsilon$; prenons ensuite un i tel que

$$|a - u_i| < \varepsilon' ;$$

alors, pour cet i,

$$|x - \varphi(u_i)| \leqslant |x - \varphi(a)| + |\varphi(a) - \varphi(u_i)| < \varepsilon + \alpha\varepsilon' = 2\varepsilon .$$

En posant $\varphi(A) = B$ on obtient de l'inclusion établie ci-dessus que

$$\lambda_N(B_\varepsilon) \leqslant r(4\varepsilon)^N \leqslant n^\ell (4\varepsilon)^N \leqslant \left(\frac{4T}{\varepsilon'}\right)^\ell (4\varepsilon)^N = C\varepsilon^{N-\ell}$$

où $C = (4T\alpha)^\ell \cdot 4^N$; ainsi,

$$\varepsilon^{-N+k}\lambda_N(B_\varepsilon) \leqslant C\varepsilon^{k-\ell} \longrightarrow 0 \quad \text{lorsque } \varepsilon \to 0$$

(car $k - \ell \geqslant 1$) ce qui montre que $B = \varphi(A)$ est k-négligeable. ♦

(v) *Si A est une partie compacte de $\mathbb{R}^N$ contenue dans une sous-variété différentielle (de classe C^1) de dimension ℓ, $1 \leqslant \ell < k$, alors A est k-négligeable.*

DÉMONSTRATION. Si A est contenu dans un morceau M de la sous-variété de dimension ℓ alors la k-négligeabilité de A est une conséquence de (iv). Le cas général s'obtient en écrivant A comme une réunion finie $A_1 \cup \cdots \cup A_n$ où chaque A_i est une partie compacte de $\mathbb{R}^N$ contenue dans un morceau de la sous-variété de dimension ℓ. ♦

(vi) *Si $A \subset \mathbb{R}^N$ est ℓ-négligeable alors A est k-négligeable si $k > \ell$.*

Ceci est une conséquence immédiate de la définition de k-négligeabilité.

7.3.10 Exercices

1 Soient $T : \mathbb{R}^N \to \mathbb{R}^N$ une transformation linéaire inversible et X une sous-variété différentielle de dimension k ($1 \leqslant k \leqslant N$) dans $\mathbb{R}^N$; posons $T(X) = Y$. Vérifier que Y est une sous-variété différentielle de dimension k ; en particulier, si T est une homothétie de rapport $t \in \mathbb{R}^*$ ($Tx = tx$, $x \in \mathbb{R}^N$), alors $\sigma^Y = |t|^k \sigma^X$.

2 Soit $f \in \mathbb{R}^N \to \mathbb{R}$ une fonction intégrable. L'objectif de cet exercice est d'établir la formule suivante :

$$\int_{\mathbb{R}^N} f(x)\,\mathrm{d}x = \int_0^\infty \left\{ \int_{rS^{N-1}} f(u)\,\mathrm{d}\sigma^r(u) \right\} \mathrm{d}r$$
$$= \int_0^\infty r^{N-1} \left\{ \int_{S^{N-1}} f(ru)\,\mathrm{d}\sigma(u) \right\} \mathrm{d}r$$

où $S^{N-1} = \{x \in \mathbb{R}^N : \|x\| = 1\}$, σ la mesure naturelle de S^{N-1}, σ^r celle de rS^{N-1}.

On peut procéder comme suit :

(i) soit $B = \{v = (v_1, \ldots, v_{N-1}) : \|v\| < 1\}$, $H_\pm = \{(x_1, \ldots, x_N) : \pm x_N > 0\}$; posons

$$\varphi(v_1, \ldots, v_{N-1}, r) = \left(rv_1, \ldots, rv_{N-1}, r\sqrt{1 - \|v\|^2}\right) ;$$

vérifier que $\varphi : B \times]0, \infty[\to H_+$ est une application bijective de classe C^1 avec

$$\det \varphi'(v, r) = r^{N-1}\left(1 - \|v\|^2\right)^{-1/2}$$

d'où

$$\int_{H_+} f(x)\,\mathrm{d}x = \int_{B \times]0,\infty[} f\left(rv\,,\, r\sqrt{1 - \|v\|^2}\right) \frac{r^{N-1}}{\sqrt{1 - \|v\|^2}}\,\mathrm{d}r\,\mathrm{d}v ;$$

(ii) soit $S_+^{N-1} = \{x = (x_1, \ldots, x_N) \in S^{N-1} : x_N > 0\}$; représenter S_+^{N-1} par l'équation $x_N = +\left(1 - x_1^2 - \cdots - x_{N-1}^2\right)^{1/2}$, $(x_1, \ldots, x_{N-1}) \in B$ et vérifier que l'élément de volume dans S_+^{N-1} est donné par

$$\mathrm{d}\sigma = \left(1 - x_1^2 - \cdots - x_{N-1}^2\right)^{-1/2} \mathrm{d}x_1 \cdots \mathrm{d}x_{N-1} ;$$

en conclure que

$$\int_{H_+} f(x)\,\mathrm{d}x = \int_0^\infty r^{N-1} \left\{ \int_{S_+^{N-1}} f(ru)\,\mathrm{d}\sigma(u) \right\} \mathrm{d}r ;$$

en raisonnant de manière analogue sur H_- et en utilisant l'exercice 1, on obtient la formule voulue.

3 Soit σ la mesure naturelle de S^{N-1}, la sphère unité dans $\mathbb{R}^N$ (voir Ex. 2). Démontrer que

$$\omega_N = \sigma(S^{N-1}) = 2\pi^{N/2}/\Gamma(N/2)$$

en procédant comme suit : poser $B(R) = \{x \in \mathbb{R}^N : \|x\| < R\}$; prendre $f = 1_{B(R)}$ dans l'exercice 2 ; on obtient

$$\lambda_N\big(B(R)\big) = \frac{R^N}{N}\,\omega_N\ ;$$

de la formule $\lambda_N\big(B(1)\big) = \pi^{N/2}/\Gamma(N/2+1)$ (Ex. 14, sect. 5.10), on obtient ω_N.

4 Soient $S_r^{N-1} = \{x \in \mathbb{R}^N : \|x\| = r\}$; si σ est la mesure naturelle de S_r^{N-1}, vérifier que

$$\sigma(S_r^{N-1}) = \omega_N r^{N-1}$$

et constater que

$$\sigma(S_r^{N-1}) = \frac{\mathrm{d}}{\mathrm{d}r}\,\lambda_N\big(B(r)\big)$$

où $B(r)$ est comme dans l'exercice 3.

5 Soit $f : \mathbb{R}^N \to \mathbb{R}$ une fonction de la forme $f(x) = g\big(\|x\|\big)$, $g : \mathbb{R}_+ \to \mathbb{R}$ mesurable ; démontrer que f est intégrable si et seulement si $r \mapsto r^{N-1}g(r)$ est intégrable sur $\mathbb{R}_+$ et alors

$$\int_{\mathbb{R}^N} f(x)\,\mathrm{d}x = \omega_N \int_0^\infty r^{N-1}g(r)\,\mathrm{d}r$$

où ω_N est comme dans l'exercice 3.

6 Soit $f : \mathbb{R}^N \to \mathbb{R}$ une fonction mesurable ; supposons que $\big|f(x)\big| \leqslant g\big(\|x\|\big)$, $g : \mathbb{R}_+ \to \mathbb{R}$ étant telle que $r \mapsto r^{N-1}g(r)$ est intégrable ; démontrer que f est intégrable. En conclure que :

(i) si f est localement intégrable et $\big|f(x)\big| \leqslant C\|x\|^{-p}$ pour un $p > N$, un $C > 0$ et x tel que $\|x\| > R$ ($R > 0$ fixe mais quelconque) alors f est intégrable ;

(ii) si f est de support compact et $\big|f(x)\big| \leqslant C\|x\|^{-p}$ pour un $p < N$, un $C > 0$ et x tel que $\|x\| < \varepsilon$ ($\varepsilon > 0$ fixe mais quelconque) alors f est intégrable.

7 Soient X, Y deux variétés différentielles de dimension k, ℓ respectivement dans $\mathbb{R}^N$. Vérifier que $X \times Y$ est une variété différentielle de dimension $k+\ell$ dans $\mathbb{R}^{2N}$ et que

$$\sigma^{X\times Y}(A \times B) = \sigma^X(A)\,\sigma^Y(B)$$

si $A \subset X$, $B \subset Y$ (A, B boréliens). Plus généralement, si $f : X \to \mathbb{R}$, $g : Y \to \mathbb{R}$ sont deux fonctions intégrables par rapport aux mesures naturelles respectives alors $fg : X \times Y \to \mathbb{R}$ l'est aussi et

$$\int_{X\times Y} fg\,\mathrm{d}\sigma^{X\times Y} = \int_X f\,\mathrm{d}\sigma^X \cdot \int_Y g\,\mathrm{d}\sigma^Y .$$

8 Soit X une sous-variété différentielle dans $\mathbb{R}^N$ de dimension k (de classe C^1). Si A est une partie compacte de X alors $\sigma^X(A) < \infty$; en particulier, si X est compact, $\sigma^X(X) < \infty$. (Considérer d'abord le cas où $A \subset M$ est un morceau.)

9 Soit $f(x) = x \sin 1/x$, si $x \neq 0$, $f(0) = 0$; posons

$$X = \left\{(x,y) : y = f(x)\,,\ 0 < x < \frac{1}{\pi}\right\} .$$

Vérifier que X est une sous-variété différentielle dans $\mathbb{R}^2$ de dimension 1 (de classe C^∞) avec $\sigma^X(X) = \infty$ même si X est un sous-ensemble borné de $\mathbb{R}^2$.

10 Vérifier en détail les affirmations suivantes :

(i) une sous-variété différentielle X de dimension N dans $\mathbb{R}^N$ (de classe C^1) est une partie ouverte de $\mathbb{R}^N$;

(ii) la mesure naturelle de X dans (i) est donnée par la mesure de Lebesgue λ_N.

11 Soit A une partie de $\mathbb{R}^N$ telle que, pour chaque $\delta > 0$, il existe n_δ cubes de côtés de longueur δ ayant une réunion contenant A ; si $n_\delta \delta^k \to 0$ lorsque $\delta \to 0$ démontrer que A est k-négligeable.

12 Un ensemble de cardinalité finie dans $\mathbb{R}^N$ est 1-négligeable ; par contre, il existe des ensembles dénombrables qui ne sont pas k-négligeables pour $k \leqslant N$ quelconque.

13 Vérifier directement à partir de la définition que l'ensemble $A = \{(x,0) : 0 \leqslant x \leqslant 1\} \subset \mathbb{R}^2$ est 2-négligeable mais non 1-négligeable.

7.4 THÉORÈME DE LA DIVERGENCE DANS $\mathbb{R}^N$

7.4.1 Introduction

Le théorème principal de cette section généralise à $\mathbb{R}^N$ le théorème de Green dans $\mathbb{R}^2$ (sect. 6.2) et le théorème de Gauss dans $\mathbb{R}^3$ (sect. 6.4). Même si le théorème concerné est un cas particulier du théorème de Stokes généralisé, nous le présentons ici séparément pour plus de clarté ; ni son énoncé, ni sa démonstration ne demandent l'utilisation des formes différentielles ; aussi, une discussion détaillée de la notion d'orientation n'est pas nécessaire pour la compréhension du théorème.

Pour la suite nous fixons les notations suivantes :

$$\mathbb{R}^+_N = \{x = (x_1, \ldots, x_N) \in \mathbb{R}^N : x_N > 0\}$$
$$\mathbb{R}^-_N = \{x = (x_1, \ldots, x_N) \in \mathbb{R}^N : x_N < 0\}$$
$$\mathbb{R}^0_N = \{x = (x_1, \ldots, x_N) \in \mathbb{R}^N : x_N = 0\} = \mathbb{R}^{N-1} \times \{0\}.$$

Si $A \subset \mathbb{R}^N$ nous écrivons

$$A^+ = A \cap \mathbb{R}^+_N, \quad A^- = A \cap \mathbb{R}^-_N, \quad A^0 = A \cap \mathbb{R}^0_N.$$

Si $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$, on écrira $x = (x'; x_N)$ avec $x' = (x_1, \ldots, x_{N-1}) \in \mathbb{R}^{N-1}$. Pour $1 \leqslant i \leqslant N$, l'écriture

$$(x_1, \ldots, \widehat{x_i}, \ldots, x_N)$$

désigne un élément de $\mathbb{R}^{N-1}$ obtenu à partir de $(x_1, \ldots, x_N)$ en supprimant la composante x_i.

7.4.2 Frontière régulière

Soit Ω une partie ouverte non vide de $\mathbb{R}^N$; on suppose que $\Omega \neq \mathbb{R}^N$ de sorte que $\operatorname{Fr}\Omega \neq \emptyset$.

Un point $p \in \operatorname{Fr}\Omega$ appartient à la ***frontière régulière*** de Ω (notée $\partial\Omega$) si la condition suivante est vérifiée : *après une permutation éventuelle des indices*, $p = (p'; p_N)$, $p' = (p_1, \ldots, p_{N-1})$, possède un voisinage ouvert du type $U' \times I$ où U' est une boule ouverte centrée en p' dans $\mathbb{R}^{N-1}$, I est un intervalle ouvert borné dans $\mathbb{R}$ contenant p_N ; de plus, il existe une fonction $h : U' \to I$ de classe C^1 telle que $h(p') = p_N$ et

$$(U' \times I) \cap \operatorname{Fr}\Omega = \{(x'; x_N) : x' \in U', \ x_N = h(x')\} \tag{7.11}$$

avec, ou bien

$$(U' \times I) \cap \Omega = \{(x'; x_N) : x' \in U', \ x_N < h(x')\} \tag{7.12}$$

ou bien

$$(U' \times I) \cap \Omega = \{(x'; x_N) : x' \in U',\ x_N > h(x')\}. \tag{7.13}$$

(Le rôle de x_N peut être joué par un autre x_i ; c'est cela que signifie « après une permutation éventuelle des indices ».)

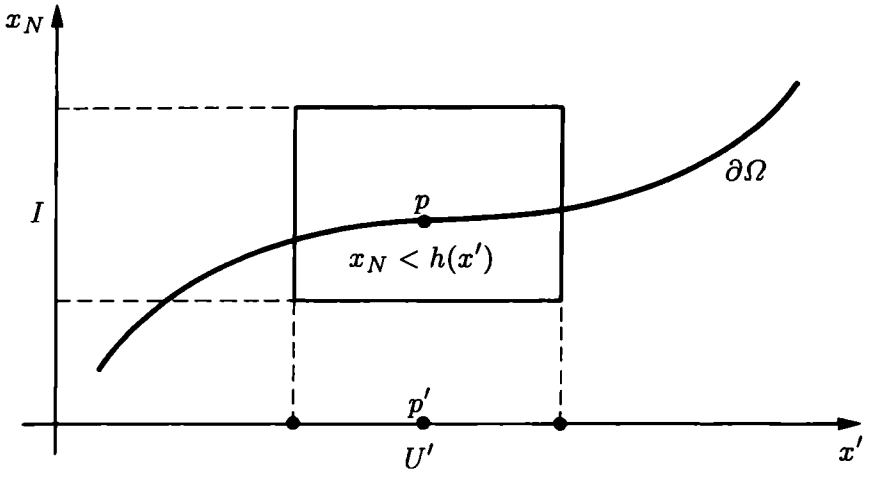


Fig. 7.2

Notons les conséquences immédiates de la définition : *si* $\partial\Omega \neq \emptyset$ *alors* $\partial\Omega$ *est une hypersurface de classe* C^1 *dans* $\mathbb{R}^N$ *et dans (7.11)*

$$(U' \times I) \cap \operatorname{Fr}\Omega = (U' \times I) \cap \partial\Omega.$$

Posons

$$S_\Omega = (\operatorname{Fr}\Omega) \setminus \partial\Omega\ ;$$

nous dirons que S_Ω est la ***frontière non régulière*** de Ω.

Donnons quelques explications géométriques. Un point p de $\operatorname{Fr}\Omega$ appartient à $\partial\Omega$ si dans un voisinage ouvert de p (du type produit $U' \times I$), $\partial\Omega$ est le graphe d'une fonction h de classe C^1 et Ω se trouve soit « au-dessous », soit « au-dessus » de ce graphe. La description analytique de $\partial\Omega$ donnée par (7.11) et (7.12) ou (7.13) est la seule qui nous sera utile dans la démonstration du théorème de la divergence. Néanmoins, essayons de clarifier la situation géométrique un peu plus. Si $p \in \operatorname{Fr}\Omega$ on peut avoir exactement une des deux situations suivantes : ou bien il existe un voisinage ouvert A de p tel que $A \cap \operatorname{Fr}\Omega$ est une hypersurface de classe C^1 ou bien pour tout voisinage ouvert A de p ce n'est pas le cas. Si l'on est dans la première situation, nous disons que $\operatorname{Fr}\Omega$ *est régulière autour de* p et nous écrivons $p \in R_\Omega$. Certainement, si $p \in \partial\Omega$ alors $p \in R_\Omega$ mais nous verrons des exemples simples dans la suite où la réciproque n'est pas vraie. Supposons que $p \in R_\Omega$; alors $A \cap \operatorname{Fr}\Omega$ est une hypersurface de classe C^1 pour un voisinage ouvert approprié A de p ; pour cette hypersurface on peut trouver une carte redressante (φ, U) dans un voisinage ouvert U de p

(prop. 4.4.4) du type suivant : $U \subset A$, $\varphi : U \to V$ est un C^1-difféomorphisme, V étant une boule ouverte centrée en $0 \in \mathbb{R}^N$, tel que $\varphi(p) = 0$ et

$$\varphi(U \cap \operatorname{Fr} \Omega) = V^0 = \{(x_1, \ldots, x_N) \in V : x_N = 0\}.$$

Il y a alors deux cas.

(i) $U \cap \operatorname{ext} \Omega$ *est non vide ; dans ce cas,* $U \setminus \operatorname{Fr} \Omega$ *aura exactement deux composantes :* $U \cap \Omega$, $U \cap \operatorname{ext} \Omega$.

En effet, puisque

$$U \setminus \operatorname{Fr} \Omega = (U \cap \Omega) \cup (U \cap \operatorname{ext} \Omega)$$

où $U \cap \Omega$ (nécessairement non vide) et $U \cap \operatorname{ext} \Omega$ sont deux ensembles ouverts disjoints non vides, on voit que $U \setminus \operatorname{Fr} \Omega$ a au moins deux composantes ; or, $V \setminus V^0$ possède exactement deux composantes V^+, V^- et $\varphi : U \to V$ est un homéomorphisme ; donc on aura ou bien

$$\varphi(U \cap \Omega) = V^+, \quad \varphi(U \cap \operatorname{ext} \Omega) = V^-$$

ou bien

$$\varphi(U \cap \Omega) = V^-, \quad \varphi(U \cap \operatorname{ext} \Omega) = V^+.$$

(Par ailleurs, en changeant φ à $-\varphi$ si besoin, on peut toujours supposer que $\varphi(U \cap \Omega) = V^-$.) Ceci établit notre affirmation.

(ii) $U \cap \operatorname{ext} \Omega$ *est vide ; dans ce cas on aura*

$$\varphi(U \cap \Omega) = V^+ \cup V^-.$$

Les points p de R_Ω dans le cas (ii) ne peuvent pas être dans $\partial\Omega$. Continuons notre analyse en nous servant des notations $\varphi : U \to V$, $\varphi(p) = 0$, introduites ci-dessus. Posons $\varphi = (\varphi_1, \ldots, \varphi_N)$; φ étant un C^1-difféomorphisme, $\operatorname{rang} \varphi'(x) = N$, $x \in U$; écrivons $\theta(x) = \varphi_N(x)$, $x \in U$. Alors $\theta : U \to \mathbb{R}$ est de classe C^1 avec $\operatorname{grad} \theta(x) \neq 0$ pour $x \in U$; si l'on est dans le cas (i) ci-dessus, on aura

$$U \cap \operatorname{Fr} \Omega = \{x \in U : \theta(x) = 0\}, \quad U \cap \Omega = \{x \in U : \theta(x) < 0\} \tag{7.14}$$

et par conséquent

$$U \cap (\operatorname{ext} \Omega) = \{x \in U : \theta(x) > 0\}.$$

En revenant à notre définition du départ de $\partial\Omega$ on voit que ces conditions sur θ seront remplies si l'on pose

$$U = U' \times I, \quad \theta(x) = x_N - h(x') \text{ ou } h(x') - x_N \quad \text{avec } x = (x'; x_N).$$

D'autre part, si $p \in \operatorname{Fr} \Omega$ possède un voisinage ouvert U et une fonction $\theta : U \to \mathbb{R}$ de classe C^1, $\operatorname{grad} \theta(x) \neq 0$ pour $x \in U$ tels que (7.14) est satisfait alors

$p \in \partial\Omega$ dans le sens de notre définition de départ utilisant (7.11), (7.12) et (7.13). Nous ne donnons pas une démonstration détaillée ; c'est une conséquence du théorème des fonctions implicites appliqué à l'équation $\theta(x) = 0$.

Les points de $\operatorname{Fr}\Omega \setminus \partial\Omega$ sont de deux types :

(a) les points qui sont dans $R_\Omega \setminus \partial\Omega$; $\operatorname{Fr}\Omega$ est régulière autour de ces points et $\operatorname{ext}\Omega$ est absent dans certains voisinages ouverts de ces points ;

(b) les points qui sont dans $\operatorname{Fr}\Omega \setminus R_\Omega$; ce sont les points où $\operatorname{Fr}\Omega$ est d'une « dimension » inférieure à $N-1$ ou ce sont des points « anguleux » de $\operatorname{Fr}\Omega$.

Soit $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \neq 0\}$; ici

$$\operatorname{Fr}\Omega = R_\Omega = \text{l'axe des } x_1\,, \quad \partial\Omega = \emptyset\,, \quad S_\Omega = \operatorname{Fr}\Omega \setminus \partial\Omega = R_\Omega\,.$$

Voici un autre exemple simple. Dans $\mathbb{R}^2$, soit

$$\Omega = B(0;1) \setminus L\,, \quad L = \{(t,0) : 0 \leqslant t \leqslant 1\}$$

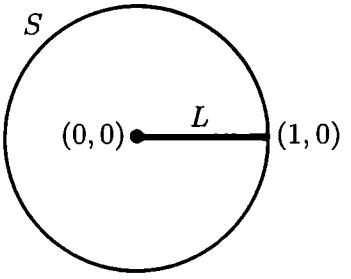


Fig. 7.3

Ici

$$\operatorname{Fr}\Omega = S \cup L\,, \quad S = S(0;1) = \{x \in \mathbb{R}^2 : \|x\| = 1\}\,;$$
$$R_\Omega = [S \setminus \{(1,0)\}] \cup [L \setminus \{(0,0)\,, (1,0)\}]$$
$$\partial\Omega = S \setminus \{(1,0)\}$$
$$S_\Omega = \operatorname{Fr}\Omega \setminus \partial\Omega = L\,.$$

De nos définitions on a également

$$\operatorname{Fr}\Omega \supset R_\Omega \supset \partial\Omega\,, \quad S_\Omega = \operatorname{Fr}\Omega \setminus \partial\Omega$$

où Ω est une partie ouverte non vide de $\mathbb{R}^N$. Si Ω est de plus bornée, on aura $\operatorname{Fr}\Omega$ compact, $\partial\Omega$ ouvert dans $\operatorname{Fr}\Omega$ et S_Ω compact.

7.4.3 Normale dirigée vers l'extérieur

On maintient les notations du paragraphe 7.4.2. Soit Ω une partie ouverte non vide de $\mathbb{R}^N$; si $p \in \partial\Omega$, on a un voisinage ouvert U de p et une fonction de classe C^1, $\theta : U \to \mathbb{R}$, tels que

$$\operatorname{grad}\theta(x) \neq 0\,, \quad x \in U$$

$$U \cap \operatorname{Fr} \Omega = \{x \in U : \theta(x) = 0\}$$
$$U \cap \Omega = \{x \in U : \theta(x) < 0\}$$
$$U \cap \operatorname{ext} \Omega = \{x \in U : \theta(x) > 0\}.$$

Rappelons que $\theta = \varphi_N$, $\varphi = (\varphi_1, \ldots, \varphi_N)$, (φ, U) étant une carte redressante standard appropriée au point p pour l'hypersurface $\partial\Omega$ et que

$$U \cap \partial\Omega = U \cap \operatorname{Fr} \Omega.$$

Posons

$$\nu(p) = \frac{\nabla\theta(p)}{\|\nabla\theta(p)\|};$$

en écrivant ν pour $\nu(p)$ on voit que, pour $|t|$ suffisamment petit, on a

$$\theta(p + t\nu) = t\,\nabla\theta(\widetilde{p}) \cdot \nu$$

où $\widetilde{p} \in\,]p, p + t\nu[$ (car $\theta(p) = 0$) ; θ étant de classe C^1,

$$\frac{\nabla\theta(\widetilde{p})}{\|\nabla\theta(\widetilde{p})\|} \cdot \nu \longrightarrow \nu \cdot \nu = 1$$

lorsque $t \to 0$. On en conclut que, pour $\varepsilon > 0$ approprié,

$$\theta(p + t\nu) > 0, \quad \theta(p - t\nu) < 0, \quad \text{si } 0 < t < \varepsilon.$$

Donc,

$$p + t\nu(p) \in U \cap \operatorname{ext} \Omega, \quad p - t\nu(p) \in U \cap \Omega, \quad 0 < t < \varepsilon. \tag{7.15}$$

De plus, $\nu = \nu(p) \in T_p^{\perp}$ où T_p est l'espace tangent à l'hypersurface $\partial\Omega$ au point p ; en effet, si γ est un chemin de classe C^1 dans $U \cap \partial\Omega$,

$$\gamma : [-1, 1] \longrightarrow U \cap \partial\Omega, \quad \gamma(0) = p,$$

alors $\theta \circ \gamma = 0$ d'où

$$\nabla\theta(p) \cdot \gamma'(0) = 0.$$

Une normale unitaire $\nu(p) \in T_p^{\perp}$ satisfaisant à la propriété

$$p + t\nu(p) \in \operatorname{ext} \Omega, \quad 0 < t < \varepsilon,$$

sera appelée la normale unitaire (à $\partial\Omega$ en p) ***dirigée vers l'extérieur***. Il est clair de (7.15) que $\nu(p)$ est uniquement déterminé ; en effet, $T_p^{\perp}$ étant de dimension 1, une autre normale unitaire à $\partial\Omega$ en p sera nécessairement $-\nu(p)$ qui, d'après (7.15), a la propriété que

$$p + t\big(-\nu(p)\big) \in \Omega, \quad 0 < t < \varepsilon ;$$

on dira que $-\nu(p)$ est la normale unitaire (à $\partial\Omega$ en p) ***dirigée vers l'intérieur***.

Lorsque le voisinage ouvert du point $p \in \partial\Omega$ est précisé par (7.11) et (7.12) (resp. (7.13)) on choisit la fonction θ ci-dessus comme

$$\theta(x'; x_N) = x_N - h(x') \quad (\text{resp. } \theta(x'; x_N) = h(x') - x_N)\,.$$

La normale unitaire dirigée vers l'extérieur sera alors donnée par

$$\nu(p) = \frac{(-\nabla h(p'); 1)}{\left\{1 + \|\nabla h(p')\|^2\right\}^{1/2}}\,, \quad \text{resp. } \nu(p) = \frac{(\nabla h(p'); -1)}{\left\{1 + \|\nabla h(p')\|^2\right\}^{1/2}}\,, \tag{7.16}$$

où $p = (p'; p_N)$, $p' = (p_1, \ldots, p_{N-1})$. Lorsque $p \notin \partial\Omega$, la définition est impossible ; ainsi, $\nu(p)$ est définie seulement pour $p \in \partial\Omega$; notons que ν est une fonction continue sur $\partial\Omega$ (car θ est de classe C^1 dans la discussion précédente). Nous donnons maintenant deux propositions qui nous serviront pour la démonstration du théorème de la divergence.

Si $F = (F_1, \ldots, F_N)$ est un champ vectoriel dérivable dans une partie ouverte de $\mathbb{R}^N$, on définit $\operatorname{div} F$ (***divergence*** de F) par la formule

$$\operatorname{div} F = \sum_{i=1}^{N} \mathrm{D}_i F_i\,.$$

7.4.4 Proposition

Soit U une partie ouverte de $\mathbb{R}^n$; alors pour $f \in C^1(U)$ tel que $\operatorname{supp} f$ *est un ensemble compact contenu dans U on a*

$$\int_U \mathrm{D}_i f \, \mathrm{d}x = 0\,, \quad i = 1, \ldots, N\,.$$

Si $F \in C^1(U; \mathbb{R}^N)$ avec $\operatorname{supp} F$ *compact contenu dans U alors*

$$\int_U (\operatorname{div} F)\, \mathrm{d}x = 0\,.$$

DÉMONSTRATION. En posant $f(x) = 0$ si $x \in U^c$ on peut supposer que $f \in C^1(\mathbb{R}^N)$; soit $T > 0$ tel que

$$\operatorname{supp} f \subset \left[-T\,, T\right]^N\,;$$

alors, si $x = (x_1, \ldots, x_N)$, on aura $f(x) = 0$ dès que $|x_i| \geqslant T$ pour un indice i. Donc,

$$\begin{aligned}\int_{-T}^{T} \mathrm{D}_i f(x_1, \ldots, x_i, \ldots, x_N)\, \mathrm{d}x_i \\ = f(x_1, \ldots, T, \ldots, x_N) - f(x_1, \ldots, -T, \ldots, x_N) = 0\,;\end{aligned}$$

d'où, d'après le théorème de Fubini,

$$\int_{\mathbb{R}^N} f(x)\,dx = \int_{-T}^{T} \overset{(N-1)}{\cdots} \int_{-T}^{T} dx_1 \cdots \widehat{dx_i} \cdots dx_N \int_{-T}^{T} D_i f\,dx_i = 0\,.$$

Aussi, si $F = (F_1, \ldots, F_N)$,

$$\int_U \operatorname{div} F\,dx = \sum_{i=1}^{N} \int_U D_i F_i\,dx = 0$$

car chaque F_i est à support compact dans U. ♦

7.4.5 Proposition

Soient U' une partie ouverte non vide de $\mathbb{R}^{N-1}$, $I =]a, b[$ un intervalle ouvert borné dans $\mathbb{R}$ et $h : U' \to I$ une fonction de classe C^1. Posons

$$A_1 = \{(x'; x_N) \in U' \times I : x_N < h(x')\}$$
$$A_2 = \{(x'; x_N) \in U' \times I : x_N > h(x')\}$$
$$H = \{(x'; x_N) \in U' \times I : x_N = h(x')\}\,.$$

Si $f : U' \times I \to \mathbb{R}$ est une fonction de classe C^1 ayant un support compact contenu dans $U' \times I$ alors

$$\int_{A_j} D_i f\,dx = (-1)^{j-1} \int_H f(x)\,\nu_i(x)\,d\sigma(x)\,, \quad i = 1, \ldots, N\,,\ j = 1, 2\,, \tag{7.17}$$

où $\nu(x) = (\nu_1(x), \ldots, \nu_N(x)) = (-\nabla h(x'); 1)/\{1 + \|\nabla h(x')\|^2\}^{1/2}$ et σ est la mesure naturelle σ^H de l'hypersurface H de classe C^1.

Si $F \in C^1(U' \times I; \mathbb{R}^N)$ avec $\operatorname{supp} F$ contenu dans $U' \times I$ alors

$$\int_{A_j} \operatorname{div} F\,dx = (-1)^{j-1} \int_H F(x) \cdot \nu(x)\,d\sigma(x)\,, \quad j = 1, 2\,. \tag{7.18}$$

REMARQUE. Notons que A_j ($j = 1$ ou 2) est une partie ouverte de $\mathbb{R}^N$ et $\partial A_j = H$; aussi, $\nu(x)$ est la normale unitaire à $H = \partial A_1$ en x dirigée vers l'extérieur (de A_1, c'est-à-dire dirigée vers l'intérieur de A_2) et $-\nu(x)$ est la normale unitaire à $H = \partial A_2$ en x dirigée vers l'extérieur (de A_2, c'est-à-dire dirigée vers l'intérieur de A_1), (§ 7.4.3, (7.16)).

DÉMONSTRATION. Prenons d'abord le cas $i = N$; on a

$$\int_a^{h(x')} D_N f(x'; x_N)\,dx_N = f(x'; h(x'))$$

$$\int_{h(x')}^{b} D_N f(x'; x_N)\,dx_N = -f(x'; h(x'))$$

car $f(x';a) = 0$, $f(x';b) = 0$ à cause du fait que supp f est une partie compacte contenue dans $U' \times]a, b[$. D'après le théorème de Fubini,

$$\int_{A_1} \mathrm{D}_N f \, \mathrm{d}x = \int_{U'} \mathrm{d}x' \int_a^{h(x')} \mathrm{D}_N f(x'; x_N) \, \mathrm{d}x_N = \int_{U'} f\big(x'; h(x')\big) \, \mathrm{d}x'$$

$$\int_{A_2} \mathrm{D}_N f \, \mathrm{d}x = \int_{U'} \mathrm{d}x' \int_{h(x')}^{b} \mathrm{D}_N f(x'; x_N) \, \mathrm{d}x_N = -\int_{U'} f\big(x'; h(x')\big) \, \mathrm{d}x' .$$

De (7.9) on a

$$\mathrm{d}\sigma(x) = \Big\{1 + \big\|\nabla h(x')\big\|^2\Big\}^{1/2} \mathrm{d}x' ;$$

donc

$$\nu_N(x) \, \mathrm{d}\sigma(x) = \mathrm{d}x'$$

d'où l'égalité

$$\int_{U'} f\big(x'; h(x')\big) \, \mathrm{d}x' = \int_H f(x) \, \nu_N(x) \, \mathrm{d}\sigma(x)$$

selon la définition du membre de droite. Ceci établit (7.17) pour $i = N$.

Soit $1 \leqslant i \leqslant N-1$; pour $x = (x'; x_N)$, $x' = (x_1, \ldots, x_{N-1})$, posons

$$\left.\begin{aligned} g_1(x) = g_1(x'; x_N) &= \int_a^{x_N} f(x'; t) \, \mathrm{d}t \\ g_2(x) = g_2(x'; x_N) &= \int_{x_N}^{b} f(x'; t) \, \mathrm{d}t \end{aligned}\right\} , \quad (x', x_N) \in U' \times I .$$

Notons que g_j est à support compact dans $U' \times I$. On a

$$\mathrm{D}_N g_1(x'; x_N) = f(x'; x_N) , \qquad \mathrm{D}_N g_2(x'; x_N) = -f(x'; x_N) ,$$

$$\mathrm{D}_i g_1(x'; x_N) = \int_a^{x_N} \mathrm{D}_i f(x'; t) \, \mathrm{d}t , \quad \mathrm{D}_i g_2(x'; x_N) = \int_{x_N}^{b} \mathrm{D}_i f(x'; t) \, \mathrm{d}t .$$

De la formule de dérivation des fonctions composées, on a

$$\begin{aligned} \mathrm{D}_i \left[g_1\big(x'; h(x')\big)\right] &= \mathrm{D}_i g_1\big(x'; h(x')\big) + \mathrm{D}_N g_1\big(x'; h(x')\big) \mathrm{D}_i h(x') \\ &= \int_a^{h(x')} \mathrm{D}_i f(x'; t) \, \mathrm{d}t + f\big(x'; h(x')\big) \mathrm{D}_i h(x') ; \end{aligned}$$

$$\begin{aligned} \mathrm{D}_i \left[g_2\big(x'; h(x')\big)\right] &= \mathrm{D}_i g_2\big(x'; h(x')\big) + \mathrm{D}_N g_2\big(x'; h(x')\big) \mathrm{D}_i h(x') \\ &= \int_{h(x')}^{b} \mathrm{D}_i f(x'; t) \, \mathrm{d}t - f\big(x'; h(x')\big) \mathrm{D}_i h(x') . \end{aligned}$$

De ces dernières formules et du théorème de Fubini on obtient

$$\begin{aligned}\int_{A_1} \mathrm{D}_i f(x)\,\mathrm{d}x &= \int_{U'} \mathrm{d}x' \int_a^{h(x')} \mathrm{D}_i f(x';x_N)\,\mathrm{d}x_N \\ &= \int_{U'} \left\{\mathrm{D}_i\left[g_1\big(x';h(x')\big)\right] - f\big(x';h(x')\big)\mathrm{D}_i h(x')\right\}\,\mathrm{d}x' \\ &= -\int_{U'} f\big(x';h(x')\big)\mathrm{D}_i h(x')\,\mathrm{d}x'\end{aligned}$$

car

$$\int_{U'} \mathrm{D}_i\left[g_j\big(x';h(x')\big)\right]\,\mathrm{d}x' = 0\,, \quad j = 1, 2\,,$$

grâce à la proposition 7.4.4 appliquée à la fonction $x' \mapsto g_j\big(x';h(x')\big)$, $x' \in U'$, qui est à support compact et de classe C^1 dans U'.

Un calcul analogue montre que

$$\int_{A_2} \mathrm{D}_i f(x)\,\mathrm{d}x = +\int_{U'} f\big(x';h(x')\big)\mathrm{D}_i h(x')\,\mathrm{d}x'\,.$$

D'après (7.9) et la définition de ν_i, on a

$$\nu_i(x)\,\mathrm{d}\sigma(x) = -\mathrm{D}_i h(x')\,\mathrm{d}x'\,;$$

donc

$$-\int_{U'} f\big(x';h(x')\big)\,\mathrm{d}x' = \int_H f(x)\,\nu_i(x)\,\mathrm{d}\sigma(x)$$

ce qui établit (7.17) pour tout i, j.

Aussi, si $F = (F_1, \ldots, F_N)$,

$$\begin{aligned}\int_{A_j} \operatorname{div} F\,\mathrm{d}x &= \sum_{i=1}^N \int_{A_j} \mathrm{D}_i F_i\,\mathrm{d}x = (-1)^{j-1}\int_H \sum_{i=1}^N F_i(x)\cdot\nu_i(x)\,\mathrm{d}\sigma(x) \\ &= (-1)^{j-1}\int_H F\cdot\nu\,\mathrm{d}\sigma\end{aligned}$$

ce qui démontre (7.18). ♦

7.4.6 Proposition (*théorème de la divergence dans* $\mathbb{R}^N$)

Soit Ω une partie ouverte non vide bornée de $\mathbb{R}^N$, $N \geqslant 2$; soient $\partial\Omega$ la frontière régulière de Ω (§ 7.4.2), ν la normale unitaire à $\partial\Omega$ dirigée vers l'extérieur et $\sigma = \sigma^{\partial\Omega}$ la mesure naturelle de l'hypersurface $\partial\Omega$; nous supposons que $\sigma(\partial\Omega) < \infty$ et que $\mathrm{Fr}\,\Omega \setminus R_\Omega$ (§ 7.4.2) est $(N-1)$-négligeable. Alors, pour tout champ vectoriel $F \in C^1(\overline{\Omega};\mathbb{R}^N)$ on a

$$\int_\Omega \operatorname{div} F\,\mathrm{d}x = \int_{\partial\Omega} F\cdot\nu\,\mathrm{d}\sigma\,. \tag{7.19}$$

REMARQUES

Rappelons que (§ 7.4.2) :

$$\operatorname{Fr}\Omega \supset R_\Omega \supset \partial\Omega\,;$$

lorsque Ω est borné, $\operatorname{Fr}\Omega$ est compact ; R_Ω, $\partial\Omega$ sont toujours ouverts dans $\operatorname{Fr}\Omega$; donc

$$S = (\operatorname{Fr}\Omega) \setminus R_\Omega$$

est compact (si Ω est borné) et $S \subset S_\Omega$. Les exemples du paragraphe 7.4.2 montrent que S peut être beaucoup plus petit que S_Ω. Nous appellerons un ensemble ouvert non vide borné Ω dans $\mathbb{R}^N$ ***gaussien*** si Ω satisfait aux conditions imposées sur Ω dans la proposition 7.4.6.

Si $S_\Omega = \emptyset$, $\operatorname{Fr}\Omega = \partial\Omega$ (Ω ouvert borné), alors $\partial\Omega$ est une hypersurface compacte de classe C^1 (non nécessairement connexe) ; dans ce cas $\sigma(\partial\Omega) < \infty$ (Ex. 8, § 7.3.10). Pour un exemple où $\sigma(\partial\Omega) = \infty$ voir l'exercice 9 du paragraphe 7.3.10.

Soit $x \in R_\Omega \setminus \partial\Omega$; il existe alors un voisinage ouvert U de x tel que $U \cap \operatorname{Fr}\Omega$ est une hypersurface de classe C^1 avec $U \cap \operatorname{ext}\Omega$ vide ; donc on peut choisir un voisinage ouvert produit de x de type $U' \times I$ (après une permutation éventuelle des indices), disjoint de S, tel que (7.11) (§ 7.4.2) est justifiée et

$$(U' \times I) \cap \Omega = \{(x'; x_N) : x' \in U',\ x_N \neq h(x')\} \tag{7.20}$$

où les notations sont comme dans (7.11). Noter qu'ici

$$(U' \times I) \cap \operatorname{Fr}\Omega \subset R_\Omega \setminus \partial\Omega\,.$$

DÉMONSTRATION. Notons d'abord qu'il suffit de démontrer (7.19) pour $F \in C^1(\mathbb{R}^N; \mathbb{R}^N)$ avec $\operatorname{supp} F$ compact. En effet, si $F \in C^1(\overline{\Omega}; \mathbb{R}^N)$, il existe une partie ouverte A de $\mathbb{R}^N$ contenant $\overline{\Omega}$ et $\widetilde{F} : A \to \mathbb{R}^N$ de classe C^1 telle que $\widetilde{F}|_{\overline{\Omega}} = F$; puisque Ω est borné, $\overline{\Omega}$ est compact ; il existe donc une partie ouverte bornée B telle que

$$\overline{\Omega} \subset B \subset A\,.$$

Soit $\theta \in C^\infty(\mathbb{R}^N)$ tel que $\theta(x) = 1$ si $x \in \overline{\Omega}$ et $\theta(x) = 0$ si $x \in B^c$ (§ 7.2.5) ; donc $\operatorname{supp}\theta = K \subset \overline{B}$ et K est compact. Posons

$$G(x) = \begin{cases} \theta(x)\widetilde{F}(x) & \text{si } x \in A \\ 0 & \text{si } x \in A^c\,; \end{cases}$$

alors $G \in C^1(\mathbb{R}^N; \mathbb{R}^N)$, $\operatorname{supp} G \subset K$ compact et $G(x) = F(x)$ si $x \in \overline{\Omega}$; la validité de (7.19) pour G équivaut à celle pour F. Supposons donc que $F \in C^1(\mathbb{R}^N; \mathbb{R}^N)$ avec $\operatorname{supp} F = K$ compact. Ecrivons S pour la partie compacte $\operatorname{Fr}\Omega \setminus R_\Omega$, introduite dans les remarques ci-dessus. Dans la première étape de la démonstration nous supposons que $K \cap S$ est vide.

Pour chaque $x \in K$ nous choisissons un voisinage ouvert U_x de x, *disjoint de* S, de manière suivante :

(i) si $x \in K \cap \Omega$, alors $U_x \subset \Omega$;
(ii) si $x \in K \cap \operatorname{ext} \Omega$, alors $U_x \subset \operatorname{ext} \Omega = (\overline{\Omega})^c$;
(iii) si $x \in K \cap (R_\Omega \setminus \partial\Omega)$, U_x est de la forme $U' \times I$ du type spécifié par (7.11) et (7.20) ;
(iv) si $x \in K \cap \partial\Omega$, U_x est de la forme $U' \times I$ telle que (7.11) et (7.12) (ou (7.13)) sont satisfaites.

Puisque $K \cap S = \emptyset$, chaque point x de K est inclus dans exactement un des quatre cas ci-dessus.

Soit $\{\theta_i\}_{i \geqslant 1}$ une partition de l'unité subordonnée à $\{U_x\}_{x \in K}$; K étant compact, il existe un entier $n \geqslant 1$ et une partie ouverte V avec $K \subset V$ tels que

$$\theta_1(x) + \cdots + \theta_n(x) = 1\,, \quad x \in V$$

(prop. 7.2.2 (ii)). Supposons que $\operatorname{supp} \theta_i \subset U_{x_i}$, $i = 1, \ldots, n$; vérifions que, pour tout $i = 1, \ldots, n$,

$$\int_\Omega \operatorname{div}(\theta_i F)\, \mathrm{d}x = \int_{\partial\Omega} \theta_i F \cdot \nu \, \mathrm{d}\sigma\,. \tag{7.21}$$

En effet, si θ_i correspond à U_{x_i} de type (i) le membre de gauche est nul à cause de la proposition 7.4.4 et celui de droite est nul car $\theta_i F = 0$ sur $\operatorname{Fr} \Omega$. Si θ_i correspond à U_{x_i} de type (ii), alors $\theta_i F = 0$ sur $\overline{\Omega}$ ce qui donne la nullité des deux membres de (7.21). Si θ_i correspond à U_{x_i} de type (iii), les deux membres de (7.21) sont de nouveau nuls ; le membre de droite est nul, car dans ce cas aussi $\theta_i F = 0$ sur $\partial\Omega$, et celui de gauche vaut

$$\int_{U' \times I} \operatorname{div}(\theta_i F)\, \mathrm{d}x = \int_{A_1} \operatorname{div}(\theta_i F)\, \mathrm{d}x + \int_{A_2} \operatorname{div}(\theta_i F)\, \mathrm{d}x = 0$$

où A_1, A_2 sont comme dans la proposition 7.4.5. Finalement, si θ_i correspond à U_{x_i} de type (iv) on aura égalité dans (7.21) à cause de (7.18) appliquée à A_1 où A_2 ; voir la remarque après la proposition 7.4.5.

Notons que

$$\sum_{i=1}^n \theta_i F = F \quad \text{sur } \mathbb{R}^N$$

car si $x \in V$, $\sum_{i=1}^n \theta_i(x) = 1$ et si $x \in V^c$, $F(x) = 0$ étant donné que $\operatorname{supp} F = K \subset V$. Donc

$$\sum_{i=1}^n \int_\Omega \operatorname{div}(\theta_i F)\, \mathrm{d}x = \int_\Omega \operatorname{div} \left(\sum_{i=1}^n \theta_i F \right) \mathrm{d}x = \int_\Omega \operatorname{div} F \, \mathrm{d}x$$

et

$$\sum_{i=1}^n \int_{\partial\Omega} \theta_i F \cdot \nu \, \mathrm{d}\sigma = \int_{\partial\Omega} \sum_{i=1}^n \theta_i F \cdot \nu \, \mathrm{d}\sigma = \int_{\partial\Omega} F \cdot \nu \, \mathrm{d}\sigma$$

d'où, en faisant la somme sur $i = 1, \ldots, n$ dans (7.21), on obtient (7.19) dans le cas $K \cap S = \emptyset$.

Passons maintenant au cas où $K \cap S \neq \emptyset$. Pour chaque $\delta > 0$, le corollaire de la proposition 7.2.8 donne une fonction $\varphi_\delta \in C^\infty(\mathbb{R}^N)$ telle que $0 \leqslant \varphi_\delta \leqslant 1$ et

$$\varphi_\delta(x) = \begin{cases} 0 & \text{si } d(x,S) < \dfrac{\delta}{2} \\ 1 & \text{si } d(x,S) > \delta \end{cases}$$

avec $\left|\mathrm{D}_j \varphi_\delta(x)\right| \leqslant 2M/\delta$, $x \in \mathbb{R}^N$, M étant une constante positive indépendante de δ. Alors

$$\operatorname{supp}(\varphi_\delta F) \subset K \cap \left\{x : d(x,S) \geqslant \frac{\delta}{2}\right\} \subset K \setminus S$$

d'où en appliquant (7.19) à $\varphi_\delta F$ on a

$$\int_\Omega \operatorname{div}(\varphi_\delta F)\,\mathrm{d}x = \int_{\partial\Omega} \varphi_\delta F \cdot \nu \,\mathrm{d}\sigma\,. \tag{7.22}$$

Lorsque $\delta \to 0$, on a

$$\lim_{\delta\to 0} \varphi_\delta(x) = \begin{cases} 0 & \text{si } x \in S \\ 1 & \text{si } x \in S^c, \end{cases}$$

d'où

$$\lim_{\delta\to 0} \int_{\partial\Omega} \varphi_\delta F \cdot \nu \,\mathrm{d}\sigma = \int_{\partial\Omega} F \cdot \nu \,\mathrm{d}\sigma\,. \tag{7.23}$$

En effet, $\partial\Omega \subset S^c$, $\sigma(\partial\Omega) < \infty$ et sur $\partial\Omega$ on a

$$\left\|\varphi_\delta F \cdot \nu\right\| \leqslant \varphi_\delta \left\|F\right\| \leqslant \left\|F\right\| \leqslant C < \infty \quad \text{où } C = \sup_{x\in\overline{\Omega}} \left\{\left\|F(x)\right\| + \left|\operatorname{div} F(x)\right|\right\}\,.$$

Aussi

$$\int_\Omega \operatorname{div}(\varphi_\delta F)\,\mathrm{d}x = \int_\Omega \varphi_\delta \operatorname{div} F \,\mathrm{d}x + \int_\Omega \nabla\varphi_\delta \cdot F \,\mathrm{d}x \tag{7.24}$$

en utilisant l'identité élémentaire

$$\operatorname{div}(\varphi F) = \varphi \operatorname{div} F + \nabla\varphi \cdot F\,;$$

puisque $\Omega \subset S^c$, on a sur Ω

$$\lim_{\delta\to 0} \varphi_\delta \operatorname{div} F = \operatorname{div} F \quad \text{et} \quad \left|\varphi_\delta \operatorname{div} F\right| \leqslant \left|\operatorname{div} F\right| \leqslant C\,,$$

d'où

$$\lim_{\delta\to 0} \int_\Omega \varphi_\delta \operatorname{div} F \,\mathrm{d}x = \int_\Omega \operatorname{div} F \,\mathrm{d}x$$

car $\lambda_N(\Omega) < \infty$, Ω étant borné. Finalement,

$$\left|\int_\Omega \nabla\varphi_\delta \cdot F \,\mathrm{d}x\right| \leqslant \int_\Omega \left\|\nabla\varphi_\delta\right\| \cdot \left\|F\right\| \mathrm{d}x \leqslant C \int_\Omega \left\|\nabla\varphi_\delta\right\| \mathrm{d}x$$

et

$$\begin{aligned}\int_{\Omega}\|\nabla\varphi_\delta\|\,\mathrm{d}x &= \int_{\Omega\cap\{x:\delta/2\leqslant d(x,S)\leqslant\delta\}}\|\nabla\varphi_\delta\|\,\mathrm{d}x\\ &\leqslant \int_{\{x:d(x,S)\leqslant\delta\}}\|\nabla\varphi_\delta\|\,\mathrm{d}x\\ &\leqslant \frac{2M\sqrt{N}}{\delta}\cdot\lambda_N\left(\{x:d(x,S)\leqslant\delta\}\right)\\ &\longrightarrow 0 \text{ lorsque } \delta\to 0\end{aligned}$$

car S est $(N-1)$-négligeable (§ 7.3.9).

En vue de (7.23) et (7.24), un passage à la limite lorsque $\delta\to 0$ dans (7.22) donne (7.19) généralement. ♦

7.4.7 Remarques

Soit $\Omega =]a, b[$, un intervalle ouvert borné dans $\mathbb{R}$, et $F\in C^1([a,b])$ alors on sait du théorème fondamental du calcul différentiel et intégral que

$$\int_a^b F'(x)\,\mathrm{d}x = F(b)-F(a)\,.$$

L'analogie avec (7.19) est évidente ; ici $\partial\Omega=\{a,b\}$ et il suffit de définir l'intégrale sur $\partial\Omega$ comme une somme et prendre $\nu(a)=-1$, $\nu(b)=1$ pour incorporer ce cas de $N=1$ dans (7.19) lorsque Ω est un intervalle ouvert ; le cas de Ω, un ensemble ouvert borné dans $\mathbb{R}$ peut être inclus aussi (Ex. 11, § 7.4.8).

Il est possible de généraliser (7.19) au cas où (i) Ω n'est pas borné et (ii) F n'est pas nécessairement de classe C^1 dans $\overline{\Omega}$. La situation générale présente des difficultés analytiques considérables, déjà dans $\mathbb{R}$. Les théorèmes très généraux sont donnés dans Federer [18] mais leur lecture est difficile.

Notre démonstration montre aussi la validité de (7.19) sous la condition

$$\int_{\partial\Omega}\|F\|\,\mathrm{d}\sigma<\infty\,,$$

$F\in C^1(\overline{\Omega};\mathbb{R}^N)$. En fait, le théorème de la divergence peut être généralisé très loin ; les travaux des années 1950-1960 de Federer (1958) et de de Giorgi (1954) ont établi que si Ω est une partie ouverte bornée non vide de $\mathbb{R}^N$ alors pour tout $F\in C^1(\overline{\Omega};\mathbb{R}^N)$ on a

$$\int_{\Omega}\operatorname{div}F\,\mathrm{d}x=\int_{\operatorname{Fr}\Omega}F\cdot\nu\,\mathrm{d}\alpha_{N-1}$$

si et seulement si $\alpha_{N-1}(\operatorname{Fr}\Omega)<\infty$ où α_{N-1} est la mesure de Hausdorff $(N-1)$-dimensionnelle et $\nu:\operatorname{Fr}\Omega\to\mathbb{R}^N$, $\|\nu(x)\|=1$, α_{N-1}-p.p.(x), ν jouant le rôle d'une normale unitaire dirigée vers l'extérieur. Nous ne définirons pas α_{N-1} ici sauf pour dire que si $\operatorname{Fr}\Omega$ est une hypersurface compacte de classe C^1,

α_{N-1} est la mesure naturelle de celle-ci. Les références exactes aux travaux de Federer et de de Giorgi sont données dans Federer [18]. Les recherches récentes des dernières années sur le théorème de la divergence ont atteint des résultats qui semblent être presque définitifs.

Notons que l'on peut très bien avoir un domaine borné Ω tel que $\partial\Omega = \emptyset$. Il suffit de prendre une courbe de Jordan C dans $\mathbb{R}^2$ qui ne possède pas de tangente en aucun de ses points et définir Ω comme le domaine intérieur donné par C ; pour une telle courbe C, on peut démontrer que $\alpha_1(C)$ est ∞ de sorte que le théorème de la divergence, même dans sa forme générale due à Federer et de de Giorgi, est inapplicable. Les courbes irrégulières C de ce type sont très étudiées ces dernières années dans le cadre des ensembles fractals.

L'article de H. König « *Ein einfacher Beweis des Gaußchen Integralsatzes* », (*Jahresbericht der DMV*, Vol. 66 (1964), pp. 119-138) donne une version simple du théorème de la divergence qui est assez générale et utile pour les applications ; cet article nous a suggéré en particulier l'idée physiquement évidente que la partie $R' = R_\Omega \setminus \partial\Omega$ de $\mathrm{Fr}\,\Omega$ ne joue aucun rôle dans le théorème, s'il s'agit des champs F de classe C^1 partout dans $\overline{\Omega}$. En effet, physiquement, le flux à travers R' est nul car autant de fluide passe au travers R' dans un sens que dans l'autre ; notre démonstration (concernant les voisinages de type (iii) ci-dessus) rend ce raisonnement précis. Nous avons profité aussi de la présentation claire de Forster [20] du théorème de la divergence et du théorème de Stokes, même si celle-ci envisage seulement le cas régulier ($\mathrm{Fr}\,\Omega = \partial\Omega$).

Notons que les domaines réguliers dans $\mathbb{R}^2$ et $\mathbb{R}^3$, pour lesquels les théorèmes de Green et de Gauss étaient donnés dans le chapitre 6, sont des cas particuliers des ensembles ouverts gaussiens envisagés dans la proposition 7.4.6.

Les formules de Green donnés dans le paragraphe 6.4.6 et les propriétés des fonctions harmoniques présentées dans le paragraphe 6.4.7 se généralisent maintenant sans peine à $\mathbb{R}^N$, $N \geqslant 2$, en utilisant la proposition 7.4.6. Nous les résumons dans les exercices du paragraphe 7.4.8 ; les raisonnements donnés dans les paragraphes 6.4.6 et 6.4.7 sont applicables sans changements, quitte à utiliser la version du théorème de la divergence donnée ici.

7.4.8 Exercices

1 Pour les ensembles ouverts Ω suivants déterminer S_Ω, R_Ω, $\partial\Omega$:

(i) $\Omega = \mathbb{R}^N \setminus \{0\}$;
(ii) $\Omega = \mathbb{R}^N \setminus \{(x_1, \ldots, x_N) : x_N = 0\}$;
(iii) $\Omega = B(0;1) \setminus \{x \in \mathbb{R}^N : \|x\| = 1/2\}$;
(iv) $\Omega = \{(x,y) : 1 > y > f(x),\ 0 < x < 1/\pi\}$ avec $f(x) = x \sin 1/x$;
(v) $\Omega = \{(x,y) : 2 > y > f(x),\ 0 < x < 1/\pi\}$ avec $f(x) = \sin 1/x$;
(vi) $\Omega = \{(x,y) : 1 > y > f(x),\ 0 < x < 1/\pi\}$ avec $f(x) = \sin 1/x$.

Pour quels Ω peut-on utiliser la proposition 7.4.6 ?

2 Soit Ω le domaine borné de l'exercice 1(iv) . Vérifier que (7.19) est juste pour $F = (P, 0)$, $P \in C^1(\overline{\Omega})$ même si $\sigma(\partial\Omega) = \infty$ ici (cf. Ex. 9, § 7.3.10).

Par la suite les symboles ν, σ et ∂ auront la signification qu'ils ont à la proposition 7.4.6.

3 Soit Ω un ensemble ouvert gaussien dans $\mathbb{R}^N$; vérifier que

$$\lambda_N(\Omega) = \frac{1}{N} \int_{\partial\Omega} x \cdot \nu(x) \, d\sigma(x) .$$

En prenant $\Omega = B(a; r)$ obtenir que

$$\lambda_N\big(B(a;r)\big) = \frac{r}{N} \sigma\big(S(a;r)\big)$$

où σ est la mesure naturelle de la sphère $S(a; r)$.

4 Soit $F \in C^1(\Omega; \mathbb{R}^N)$, Ω étant un ensemble ouvert dans $\mathbb{R}^N$. Alors

$$\operatorname{div} F(x) = \lim_{n\to\infty} \frac{1}{\lambda_N(A_n)} \int_{\partial A_n} F \cdot \nu \, d\sigma$$

où $\{A_n\}$ est une suite d'ensembles gaussiens dans Ω telle que $x \in \overline{A}_n$, $n = 1, 2, \ldots$, et diamètre$(A_n) \to 0$ $(n \to \infty)$.

5 Soit $f \in C^2(\Omega)$, Ω étant un ensemble ouvert dans $\mathbb{R}^N$. Alors

$$\Delta f(x) = \lim_{n\to\infty} \frac{1}{\lambda_N(A_n)} \int_{\partial A_n} \mathrm{D}_\nu f \, d\sigma$$

où $\{A_n\}$ est comme dans l'exercice 4. (Noter que $\Delta f = \operatorname{div} \operatorname{grad} f$ et $\mathrm{D}_\nu f = \nabla f \cdot \nu$.)

6 Soit Ω un ensemble ouvert gaussien dans $\mathbb{R}^N$. Si $f, g \in C^2(\overline{\Omega})$ démontrer les formules (de Green) suivantes (cf. § 6.4.6) :

(i) $\displaystyle\int_{\partial\Omega} f \mathrm{D}_\nu g \, d\sigma = \int_\Omega (f \, \Delta g + \nabla f \cdot \nabla g) \, dx$;

(ii) $\displaystyle\int_{\partial\Omega} (f \mathrm{D}_\nu g - g \, \mathrm{D}_\nu f) \, d\sigma = \int_\Omega (f \, \Delta g - g \, \Delta f) \, dx$;

(iii) $\displaystyle\int_{\partial\Omega} \mathrm{D}_\nu f \, d\sigma = \int_\Omega \Delta f \, dx$;

(iv) $\displaystyle\int_{\partial\Omega} \mathrm{D}_\nu f \, d\sigma = 0$ si f est harmonique dans Ω ;

(v) $\displaystyle\int_{\partial\Omega} f \, \mathrm{D}_\nu f \, d\sigma = \int_\Omega \|\nabla f\|^2 \, dx$ si f est harmonique dans Ω.

7 Soit f une fonction harmonique dans un domaine D de $\mathbb{R}^N$. Si $\overline{B}(a;r) \subset D$ alors

$$f(a) = \frac{1}{\sigma(S)} \int_S f \, d\sigma$$

où $S = S(a;r) = \{x : \|x - a\| = r\}$ (cf. § 6.4.7 ; noter que $\sigma(S) = \omega_N r^{N-1}$ est calculé dans l'exercice 3 du paragraphe 7.3.10).

8 (Suite de l'exercice 7) On a :

$$f(a) = \frac{1}{\lambda_N(B)} \int_B f(x) \, dx$$

où $B = B(a;r)$ (cf. (6.31) et Ex. 2, § 7.3.10).

9 Généraliser le principe du maximum pour les fonctions harmoniques dans $\mathbb{R}^N$ en suivant la discussion donnée dans le paragraphe 6.4.7.

10 Soit Ω un ensemble ouvert non vide borné dans $\mathbb{R}^N$. Supposons que $\theta : U \to \mathbb{R}$ est une fonction de classe C^1 définie dans un ensemble ouvert non vide U telle que $\operatorname{grad} \theta(x) \neq 0$ pour $x \in U$ avec

$$U \cap \overline{\Omega} = \{x \in U : \theta(x) \leqslant 0\}.$$

Démontrer que

$$U \cap \operatorname{Fr} \Omega = \{x \in U : \theta(x) = 0\}.$$

11 Soit Ω une partie ouverte non vide de $\mathbb{R}$. On peut démontrer que $\Omega = \bigcup_{n \geqslant 1} I_n$ où les I_n sont des intervalles ouverts deux à deux disjoints. On peut interpréter $\partial\Omega$ comme l'ensemble des points $p \in \operatorname{Fr} \Omega$ tels que :

(i) $]p - \varepsilon, p] \cap \Omega = \emptyset$ ou,

(ii) $[p, p + \varepsilon[\cap \Omega = \emptyset$ pour un $\varepsilon > 0$;

posons $\nu(p) = -1$ si p de $\operatorname{Fr} \Omega$ satisfait à (i) et $\nu(p) = 1$ si p de $\operatorname{Fr} \Omega$ satisfait (ii). On interprète $\sigma^{\partial\Omega} = \sigma$ comme la mesure de comptage de sorte que $\sigma(\partial\Omega) < \infty$ veut dire que $\partial\Omega$ contient un nombre fini de points. Avec ces conventions, vérifier (7.19) pour $F \in C^1(\overline{\Omega})$, Ω borné, au moins lorsque Ω est une réunion finie des I_n.

(On peut avoir $\partial\Omega = \{0, 1\}$ mais Ω une réunion infinie des I_n contenues dans $]0, 1[$; consulter à ce sujet la construction d'un ensemble de Cantor.)

7.5 FORMES MULTILINÉAIRES ALTERNÉES

7.5.1 Introduction

Cette section introduit succinctement un chapitre de l'algèbre multilinéaire qui permet de donner une base rigoureuse pour les calculs formels indiqués dans la section 7.1.

7.5.2 Préliminaires

Soit E un espace vectoriel réel de dimension finie $N \geqslant 1$. Rappelons que $\mathcal{L}(E;F)$, F un espace vectoriel réel, représente l'espace vectoriel formé des applications linéaires de E dans F et que $\mathcal{L}_k(E;F)$, k entier $\geqslant 1$, est l'espace vectoriel formé des applications multilinéaires de

$$E^k = E \times \overset{k}{\cdots} \times E \quad \text{dans } F\,;$$

un élément $u \in \mathcal{L}_k(E;F)$ est une application

$$u : E^k \to F$$

telle que

$$a \mapsto u(x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_k)$$

est linéaire pour tout choix de i, $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k$. Lorsque $F = \mathbb{R}$ nous écrivons $\mathcal{L}_k(E)$ au lieu de $\mathcal{L}_k(E;\mathbb{R})$; les éléments de $\mathcal{L}_k(E)$ s'appellent les ***formes k-linéaires*** (génériquement, ***formes multilinéaires***) sur E. L'espace $\mathcal{L}_1(E)$ est noté généralement comme E' ; on dit que E' est l'***espace dual*** de E ; un élément $u \in E'$ s'appelle aussi une ***forme linéaire*** sur E.

Une forme k-linéaire u s'appelle ***alternée*** si

$$u(x_1, \ldots, x_k) = 0$$

dès que $x_i = x_j$ pour deux indices $i \neq j$. Posons

$$\mathcal{A}_k(E) = \{u \in \mathcal{L}_k(E) : u \text{ est alternée}\}\,;$$

on vérifie immédiatement que $\mathcal{A}_k(E)$ est un sous-espace vectoriel de $\mathcal{L}_k(E)$. On convient que $\mathcal{L}_1(E) = \mathcal{A}_1(E)$, $\mathcal{A}_0(E) = \mathbb{R}$.

On démontre qu'*une forme k-linéaire u est alternée si et seulement si u est* ***antisymétrique***, c'est-à-dire

$$u(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_k) = -u(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_k)$$

pour tout choix de $(x_1, \ldots, x_k) \in E^k$, $i < j$ dans $\{1, \ldots, k\}$. En effet, puisque les deux propriétés « alternées », « antisymétriques » concernent deux coordonnées seulement, on peut se borner à la considération des formes bilinéaires φ ;

si φ est alternée, on a pour tout a, b dans E, en utilisant la bilinéarité de φ, que

$$0 = \varphi(a+b\,,\,a+b) = \varphi(a,a) + \varphi(a,b) + \varphi(b,a) + \varphi(b,b) = \varphi(a,b) + \varphi(b,a)$$

d'où $\varphi(a,b) = -\varphi(b,a)$, c'est-à-dire φ est antisymétrique. D'autre part, si φ est antisymétrique, on a, pour tout $a \in E$,

$$\varphi(a,a) = -\varphi(a,a)$$

d'où $\varphi(a,a) = 0$, c'est-à-dire φ est alternée.

La propriété d'antisymétrie se renforce automatiquement de manière suivante : *si* $u \in \mathcal{A}_k(E)$ *alors*

$$u(x_{\pi(1)}, \ldots, x_{\pi(k)}) = \varepsilon(\pi) u(x_1, \ldots, x_k) \tag{7.25}$$

pour toute permutation π *de* $\{1, \ldots, k\}$ *et pour tout choix de* $(x_1, \ldots, x_k) \in E^k$. Ici, $\varepsilon(\pi) = 1$ si π est une permutation paire et $\varepsilon(\pi) = -1$ sinon. En effet, u antisymétrique veut dire que (7.25) est valable lorsque π est une transposition qui échange les indices i et j, $i \neq j$, laissant les autres fixes ; rappelons que $\varepsilon(\pi) = -1$ si π est une transposition. Puisque toute permutation π est le produit d'un certain nombre r de transpositions $\pi_1, \ldots, \pi_r$ (Conventions, Notations, Rappels, § 4), on obtient (7.25) par le raisonnement suivant : posons, pour $x = (x_1, \ldots, x_k)$,

$$\pi x = (x_{\pi(1)}, \ldots, x_{\pi(k)})\,;$$

alors

$$u(\pi x) = u(\pi_1 \cdots \pi_r x) = \big(-1\big) u(\pi_2 \cdots \pi_r x) = \big(-1\big)^r u(x) = \varepsilon(\pi) u(x)\,,$$

ce qui est (7.25).

7.5.3 Produit extérieur

Soit E un espace vectoriel réel de dimension finie $N \geqslant 1$. Si $u_1, \ldots, u_k$ sont k éléments de E', nous définissons

$$u = u_1 \wedge \cdots \wedge u_k : E^k \longrightarrow \mathbb{R}$$

par la formule

$$(u_1 \wedge \cdots \wedge u_k)(x_1, \ldots, x_k) = \det \begin{bmatrix} u_1(x_1) & \ldots & u_1(x_k) \\ \vdots & \ddots & \vdots \\ u_k(x_1) & \ldots & u_k(x_k) \end{bmatrix}. \tag{7.26}$$

Les propriétés des déterminants nous donnent immédiatement que $u \in \mathcal{A}_k(E)$.

Notons par $S(k;N)$, $k \leqslant N$, l'ensemble des suites strictement croissantes d'entiers de 1 à N de longueur k ; si $I \in S(k;N)$, I est de la forme

$$I = \{i_1 < \ldots < i_k\}, \quad 1 \leqslant i_j \leqslant N ;$$

la cardinalité de $S(k;N)$ est

$$\binom{N}{k} = \frac{N!}{k!\,(N-k)!} .$$

Supposons que $\{u_1, \ldots, u_N\}$ est une base de E' ; pour $I \in S(k;N)$, $I = \{i_1 < \ldots < i_k\}$, posons

$$u^I = u_{i_1} \wedge \cdots \wedge u_{i_k} .$$

Pour $k \leqslant N$, les éléments u^I, $I \in S(k;N)$, forment une base pour $\mathcal{A}_k(E)$; en particulier, la dimension de $\mathcal{A}_k(E)$, pour $k \leqslant N$, est égale à $\binom{N}{k}$; si $k > N$, $\mathcal{A}_k(E)$ est réduit à $\{0\}$.

En effet, soit $\{e_1, \ldots, e_N\}$ une base de E, duale à $\{u_1, \ldots, u_N\}$, c'est-à-dire

$$u_i(e_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon.} \end{cases}$$

Si $J = \{j_1, \ldots, j_k\}$ avec chaque j_t dans $\{1, \ldots, N\}$, posons

$$e^J = (e_{j_1}, \ldots, e_{j_k}) ;$$

si u est un élément quelconque de $\mathcal{A}_k(E)$ on aura

$$u(e^J) = 0$$

si les indices j_t ne sont pas distincts. Si $x = (x_1, \ldots, x_k) \in E^k$, et

$$x_i = \sum_{j=1}^{N} a_{ij} e_j , \quad 1 \leqslant i \leqslant k ,$$

on aura pour toute forme k-linéaire u que

$$u(x) = \sum_J a_J\, u(e^J)$$

où $J = \{j_1, \ldots, j_k\}$, $a_J = a_{1j_1} \cdots a_{kj_k}$, la somme portant sur toutes les suites d'entiers de 1 à N de longueur k. Ceci démontre déjà que pour $k > N$, $u(x) = 0$, $x \in E^k$, pour tout $u \in \mathcal{A}_k(E)$ car alors aucun J ne peut avoir les éléments $j_1, \ldots, j_k$ distincts. Si $k \leqslant N$, ce raisonnement montre que

$$u(x) = \sum_{J \in S(k;N)} a_J u(e^J) , \quad x \in E^k ,$$

pour $u \in \mathcal{A}_k(E)$; si $u = u^I$, $I = \{i_1 < \ldots < i_k\} \in S(k;N)$, alors (7.26) donne

$$u^I(e^J) = \begin{cases} 1 & \text{si } I = J \in S(k;N) \\ 0 & \text{sinon.} \end{cases}$$

Si $u \in \mathcal{A}_k(E)$, posons

$$c_I = u(e^I), \quad I \in S(k;N) ;$$

il est alors clair que

$$u = \sum_{I \in S(k;N)} c_I u^I , \tag{7.27}$$

car, de ce qui précède,

$$u^I(x) = \sum_{J \in S(k;N)} a_J u^I(e^J) = a_I$$

d'où

$$\sum_{I \in S(k;N)} c_I u^I(x) = \sum_{I \in S(k;N)} c_I a_I = \sum_{I \in S(k;N)} a_I u(e^I) = u(x)$$

pour $x \in E^k$. Ceci établit notre affirmation.

Nous donnons maintenant la proposition principale concernant la multiplication (dite ***extérieure***) entre un élément de $\mathcal{A}_p(E)$ et un autre de $\mathcal{A}_q(E)$, résultant en un élément de $\mathcal{A}_{p+q}(E)$.

Nous fixons une base $\{u_1, \ldots, u_N\}$ pour E' et notons par $\{e_1, \ldots, e_N\}$ la base de E qui est duale à $\{u_1, \ldots, u_N\}$. Si $\alpha \in \mathcal{A}_p(E)$, on sait d'après (7.27) que

$$\alpha = \sum_{I \in S(p;N)} a_I u^I$$

où $u^I = u_{i_1} \wedge \cdots \wedge u_{i_p}$ si $I = \{i_1 < \ldots < i_p\}$, $a_I \in \mathbb{R}$, u^I étant défini par (7.26) ; de même si $\beta \in \mathcal{A}_q(E)$, on a

$$\beta = \sum_{J \in S(q;N)} b_J u^J .$$

Nous définissons $\alpha \wedge \beta \in \mathcal{A}_{p+q}(E)$ par la formule

$$\alpha \wedge \beta = \sum_{\substack{I \in S(p;N) \\ J \in S(q;N)}} a_I b_J \, u^I \wedge u^J \tag{7.28}$$

où

$$u^I \wedge u^J = (u_{i_1} \wedge \cdots \wedge u_{i_p} \wedge u_{j_1} \wedge \cdots \wedge u_{j_q}) \in \mathcal{A}_{p+q}(E)$$

si $I = \{i_1 < \ldots < i_p\}$, $J = \{j_1 < \ldots < j_q\}$. Rappelons encore que $u_{i_1} \wedge \cdots \wedge u_{j_q}$ est défini par la formule générale (7.26). Il est clair que si $p + q > N$ on aura $\alpha \wedge \beta = 0$. Les propriétés du produit extérieur sont résumées dans la proposition suivante.

7.5.4 Proposition

(i) Bilinéarité : on a

$$(\alpha_1 + \alpha_2) \wedge \beta = (\alpha_1 \wedge \beta) + (\alpha_2 \wedge \beta)$$

$$\alpha \wedge (\beta_1 + \beta_2) = (\alpha \wedge \beta_1) + (\alpha \wedge \beta_2)$$

$$(t\alpha) \wedge \beta = t(\alpha \wedge \beta) = \alpha \wedge (t\beta)$$

pour $t \in \mathbb{R}$, α, α_1, α_2 *dans* $\mathcal{A}_p(E)$, β, β_1, β_2 *dans* $\mathcal{A}_q(E)$.

(ii) Anticommutativité : on a

$$\alpha \wedge \beta = (-1)^{pq}(\beta \wedge \alpha)$$

pour $\alpha \in \mathcal{A}_p(E)$, $\beta \in \mathcal{A}_q(E)$.

(iii) Associativité : si $\alpha \in \mathcal{A}_p(E)$, $\beta \in \mathcal{A}_q(E)$ *et* $\gamma \in \mathcal{A}_r(E)$ *alors*

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma) \in \mathcal{A}_{p+q+r}(E)\,.$$

DÉMONSTRATION. (i) est immédiate de (7.28). Pour (ii) il suffit, en vue de (7.28), de vérifier que

$$(u_{i_1} \wedge \cdots \wedge u_{i_p}) \wedge (u_{j_1} \wedge \cdots \wedge u_{j_q}) = (-1)^{pq}(u_{j_1} \wedge \cdots \wedge u_{j_q}) \wedge (u_{i_1} \wedge \cdots \wedge u_{i_p})\,.$$

Ceci devient évident en remontant à la définition des deux membres comme des déterminants d'ordre $(p+q)$ donnés par la formule (7.26) et par le fait que la permutation

$$\{1, \ldots, p, p+1, \ldots, p+q\} \longrightarrow \{p+1, \ldots, p+q, 1, \ldots, p\}$$

introduit pq inversions, d'où sa signature est $(-1)^{pq}$. Pour (iii), il suffit (en vue de 7.28) de vérifier que

$$(u^I \wedge u^J) \wedge u^K = u^I \wedge (u^J \wedge u^K)$$

si $I \in S(p; N)$, $J \in S(q; N)$, $K \in S(r; N)$; or ceci est évident car, d'après les définitions, les deux membres calculés en $x \in E^{p+q+r}$ d'après (7.26) résultent dans un même déterminant d'ordre $p + q + r$. ♦

7.5.5 Remarques

En vertu de l'associativité du produit extérieur, une écriture comme

$$\alpha_1 \wedge \cdots \wedge \alpha_k$$

où $\alpha_i \in \mathcal{A}_{p_i}(E)$, $i = 1, \ldots, k$, définit un unique élément de $\mathcal{A}_p(E)$, $p = p_1 + \cdots + p_k$.

A cause de l'anticommutativité du produit extérieur, on a

$$\alpha \wedge \alpha = 0$$

pour tout $\alpha \in \mathcal{A}_p(E)$, $p \geqslant 1$, p impair.

La dimension de $\mathcal{A}_N(E)$ est égale à 1, N étant la dimension de E ; en fait, de (7.26) on voit que tout élément $\alpha \in \mathcal{A}_N(E)$ est de la forme

$$\alpha = c\, u_1 \wedge \cdots \wedge u_N$$

où $c \in \mathbb{R}$ et $\{u_1, \ldots, u_N\}$ est une base de E.

Si $\alpha \in \mathcal{A}_p(E)$ on a vu dans (7.27) que l'on a

$$\alpha = \sum_{I \in S(p;N)} a_I u^I$$

où $a_I = u^I(e^I)$, $u^I = u_{i_1} \wedge \cdots \wedge u_{i_p}$, $e^I = (e_{i_1}, \ldots, e_{i_p})$ si $I = \{i_1 < \ldots < i_p\}$ et $\{u_1, \ldots, u_N\}$, $\{e_1, \ldots, e_N\}$ sont des bases duales de E' et E respectivement ; notons que le choix de la base $\{e_1, \ldots, e_N\}$ pour E détermine la base duale $\{u_1, \ldots, u_N\}$ pour E' et vice versa. Les nombres réels $\{a_I\}$, $I \in S(p;N)$, s'appellent les ***composantes*** de α dans la base $\{e_1, \ldots, e_N\}$ de E (ou dans la base $\{u_1, \ldots, u_N\}$ de E'). Il est utile de savoir comment se transforment les composantes $\{a_I\}$, lorsque l'on passe d'une base à une autre. Ces formules de transformations peuvent être obtenues, en principe, à partir de la formule suivante : *si* $v_1, \ldots, v_m$ *sont les éléments de* E' *donnés par*

$$v_i = \sum_{j=1}^{n} c_{ij} u_j\,, \quad i = 1, \ldots, m\,,$$

où $u_1, \ldots, u_n$ *sont* n *éléments quelconques de* E' *alors*

$$v_1 \wedge \cdots \wedge v_m = 0 \quad \textit{si } m > n$$

et, si $m \leqslant n$ *et* $(x_1, \ldots, x_m) \in E^m$, *on a*

$$(v_1 \wedge \cdots \wedge v_m)(x_1, \ldots, x_m) = \det\{[c_{ij}][u_i(x_j)]\} \tag{7.29}$$

où $[c_{ij}]$ *est une matrice de type* (m,n) *et* $[u_i(x_j)]$ *est une matrice de type* (n,m).

En effet, d'après (7.26)

$$(v_1 \wedge \cdots \wedge v_m)(x_1, \ldots, x_m) = \det[v_i(x_j)]\ ;$$

or

$$v_i(x_j) = \sum_{k=1}^{n} c_{ik} u_k(x_j)\,,$$

d'où résulte la relation matricielle

$$[v_i(x_j)]_{(m,m)} = [c_{ij}]_{(m,n)} \cdot [u_i(x_j)]_{(n,m)}$$

qui donne (7.29) si $m \leqslant n$; si $m > n$, (7.29) reste vrai mais d'après la formule de Binet-Cauchy (§ 7.3.5) :

$$\det[v_i(x_j)] = 0$$

pour tout choix de $(x_1, \ldots, x_m) \in E^m$, c'est-à-dire

$$v_1 \wedge \cdots \wedge v_m = 0\,.$$

La conclusion que $v_1 \wedge \cdots \wedge v_m = 0$, si $m > n$, reflète le fait que $v_1, \ldots, v_m$ sont forcément linéairement dépendants s'ils sont obtenus comme combinaison linéaire de $n < m$ éléments de E' ; ceci peut être établi directement.

Si $m = n$, (7.29) se simplifie et on obtient

$$(v_1 \wedge \cdots \wedge v_m)(x_1, \ldots, x_m) = \det[c_{ij}] \det[u_i(x_j)]$$

ce qui s'écrit comme

$$v_1 \wedge \cdots \wedge v_m = \det[c_{ij}](u_1 \wedge \cdots \wedge u_m)$$

Si $m \leqslant n$, on peut reformuler (7.29) en utilisant la formule de Binet-Cauchy ; on aura

$$v_1 \wedge \cdots \wedge v_m = \sum_{1 \leqslant j_1 < \ldots < j_m \leqslant n} \det[c_{ij_k}]\, u_{j_1} \wedge \cdots \wedge u_{j_k} \tag{7.30}$$

où $[c_{ij_k}]$ est une matrice (m, m) formée de colonnes numéros $j_1, \ldots, j_m$ de la matrice $[c_{ij}]$ du type (m, n).

Nous n'aurons pas besoin des formules plus explicites pour les transformations des composantes $\{a_I\}$ d'un $\alpha \in \mathcal{A}_p(E)$. Les seules propriétés du produit extérieur des formes antisymétriques que nous utiliserons sont celles données dans la proposition 7.5.4 et celle contenue dans la formule (7.26). Il est clair, en fait, qu'il existe un unique produit extérieur

$$\Lambda : \mathcal{A}_p(E) \times \mathcal{A}_q(E) \longrightarrow \mathcal{A}_{p+q}(E)\,, \quad p \geqslant 1\,,\ q \geqslant 1\,,$$

compatible avec (7.26) et les propriétés énumérées dans la proposition 7.5.4. Notre présentation du produit extérieur est fortement dépendante du choix d'une base ; ceci à l'avantage d'arriver au but rapidement (prop. 7.5.4). La plupart des auteurs préfèrent définir le produit extérieur par un procédé différent : si $\alpha \in \mathcal{A}_p(E)$, $\beta \in A_q(E)$, ils posent

$$(\alpha \wedge \beta)(x_1, \ldots, x_p, x_{p+1}, \ldots, x_{p+q}) =$$
$$= \frac{1}{p!\,q!} \sum_{\pi \in S_{p+q}} \varepsilon(\pi)\alpha\big(x_{\pi(1)}, \ldots, x_{\pi(p)}\big) \cdot \beta\big(x_{\pi(p+1)}, \ldots, x_{\pi(p+q)}\big)$$

où S_{p+q} est le groupe des permutations sur $\{1, \ldots, p+q\}$ et $(x_1, \ldots, x_{p+q}) \in E^{p+q}$. Cette définition à l'avantage d'être indépendante du choix d'une base ; de plus, elle est applicable dans des situations plus générales. Mais il n'est pas évident que la définition proposée donne un produit possédant les propriétés désirées ; en particulier, l'associativité du produit ainsi défini exige une démonstration soignée.

Un autre avantage d'une définition indépendante du choix d'une base est qu'elle permet une présentation élégante de la théorie des déterminants. On définit alors un déterminant comme l'unique forme N-linéaire alternée sur $\mathbb{R}^N$ (notée det) satisfaisant à la condition

$$\det(e_1, \ldots, e_N) = 1$$

où $\{e_1, \ldots, e_N\}$ est la base canonique de $\mathbb{R}^N$. Toutes les formules concernant les déterminants (en particulier, celle de Binet-Cauchy et celle de Laplace) peuvent être déduites de cette définition.

7.5.6 Compléments

Nous avons à peine effleuré une toute petite partie, directement utile pour nous, d'un vaste chapitre des mathématiques appelé le ***calcul tensoriel***. Les éléments de $\mathcal{L}_k(E)$ s'appellent, dans ce calcul, les ***tenseurs covariants*** d'ordre k (attachés à l'espace vectoriel E) et ceux de $\mathcal{A}_k(E)$ s'appellent les ***tenseurs covariants antisymétriques*** d'ordre k.

Etant donné un espace vectoriel réel E de dimension finie $\geqslant 1$, l'espace $\mathcal{T}_r^s(E)$ formé des applications multilinéaires (à valeurs réelles) sur

$$E^r \times E'^s = E \times \overset{r}{\cdots} \times E \times E' \times \overset{s}{\cdots} \times E'$$

s'appelle l'espace des ***tenseurs*** (***mixtes d'ordre*** $r+s$), ***covariants d'ordre*** r et ***contravariants d'ordre*** s. Cette terminologie provient du fait que les éléments de E sont appelés contravariants et ceux de E' covariants ; la justification pour cette terminologie est basée sur la nature des transformations subies par les composantes des éléments de E et E' lorsque l'on passe d'une base à une autre. Selon cette terminologie, $\mathcal{L}_k(E) = \mathcal{T}_k^0(E)$.

Il faut noter que certains auteurs écrivent $\bigwedge^k E'$ pour notre $\mathcal{A}_k(E)$; l'explication pour cette écriture est la suivante. Dans la théorie générale de l'algèbre multilinéaire, une construction fondamentale est celle de la formation du ***produit tensoriel*** (noté $E_1 \otimes \cdots \otimes E_n$) d'un certain nombre d'espaces vectoriels $E_1, \ldots, E_n$. Si s de ces espaces sont l'espace F et $r = n - s$ de ces espaces sont l'espace dual F', l'espace

$$F \otimes \overset{s}{\cdots} \otimes F \otimes F' \otimes \overset{r}{\cdots} \otimes F'$$

est isomorphe à $\mathcal{T}_r^s(F)$ défini ci-dessus (si F est un espace de dimension finie) ; cet isomorphisme résulte de l'isomorphisme entre $(F')'$ et F, si F est de dimension finie. On étudie alors les tenseurs symétriques et antisymétriques dans

$\mathcal{T}_0^s(F)$; l'espace des tenseurs antisymétriques dans $\mathcal{T}_0^s(F)$ est noté $\bigwedge^s F$; $\bigwedge^s F$ est isomorphe à $\bigwedge^s F'$; il en est de même pour $\mathcal{T}_0^s(F)$ et $\mathcal{T}_s^0(F')$ (si la dimension de F est finie). L'espace $\bigoplus_{s=0}^N \bigwedge^s F$ muni du produit extérieur s'appellera l'***algèbre extérieure*** de F ; ainsi, notre calcul est celui de l'algèbre extérieure de E'.

Malheureusement, d'autres auteurs écrivent $\bigwedge^s E$ pour notre $\mathcal{A}_s(E)$; aussi, la définition de $\alpha \wedge \beta$ donnée par certains auteurs introduit un facteur scalaire supplémentaire ; pour ces raisons, nous avons utilisé une notation neutre $\mathcal{A}_k(E)$, suivant en cela d'autres auteurs ayant le même souci. Néanmoins, ce qu'il faut retenir c'est que l'objectif visé par toutes ces définitions est celui d'avoir les propriétés énumérées dans la proposition 7.4.5.

Disons un mot sur le produit tensoriel des espaces vectoriels réels $E_1, \ldots, E_n$ de dimension $N_1, \ldots N_n$ respectivement ; c'est un espace vectoriel $E = E_1 \otimes \cdots \otimes E_n$ caractérisé par les propriétés suivantes :

(i) pour $x_1 \in E_1, \ldots, x_n \in E_n$ il existe un élément $x_1 \otimes \cdots \otimes x_n \in E$;
(ii) l'application $(x_1, \ldots, x_n) \mapsto x_1 \otimes \cdots \otimes x_n$ est multilinéaire ;
(iii) si $\{e_1^i, \ldots, e_{N_i}^i\}$ est une base de E_i, $1 \leqslant i \leqslant n$, alors les éléments

$$e_{\alpha_1}^1 \otimes \cdots \otimes e_{\alpha_n}^n$$

forment une base de E.

Il y a différentes constructions pour établir l'existence et l'unicité (à l'isomorphisme près) de E. Une construction rapide est celle de poser

$$E = \mathcal{L}_n(E_1', \ldots, E_n'),$$

l'espace des formes multilinéaires sur $E_1' \times \cdots \times E_n'$; c'est la voie que nous avons choisie dans cette section.

Le calcul tensoriel est très utile pour la géométrie différentielle de dimension supérieure et, par voie de conséquence, pour la relativité générale. La théorie d'élasticité et d'autres théories voisines sont souvent formulées en termes tensoriels ; d'ailleurs, le mot « tenseur » prend son origine dans la théorie de l'élasticité.

Une application linéaire $T : V \to W$ d'un espace vectoriel V dans un autre espace vectoriel W peut être considérée comme une forme bilinéaire u sur $V \times W'$ par le procédé suivant : si $x \in V$, $\xi \in W'$, posons

$$u(x, \xi) = \xi(Tx);$$

ceci démontre que l'espace vectoriel des applications linéaires $\mathcal{L}(V; W)$ est (naturellement) isomorphe à $\mathcal{L}_2(V, W')$, l'espace des formes bilinéaires dans $V \times W'$; de plus, si V, W' sont de dimensions finies, $\mathcal{L}_2(V, W')$ est isomorphe à $V' \otimes W$; on peut alors écrire que

$$\mathcal{L}(V; W) \sim \mathcal{L}_2(V, W') \sim V' \otimes W .$$

L'identification de $\mathcal{L}(V;W)$ avec $V' \otimes W$ est souvent utilisée dans la théorie de l'élasticité pour dire qu'une matrice $(3,3)$, par exemple, est un tenseur (en prenant $V = \mathbb{R}^3$, $W = \mathbb{R}^3$). Le petit livre de Lichnerowicz [24] donne une introduction intéressante au calcul tensoriel et ses applications.

7.5.7 Exercices

[1] Soient $E_1, \ldots, E_r, F, (r+1)$ espaces vectoriels réels et $u : E_1 \times \cdots \times E_r \to F$ une application multilinéaire. Vérifier par récurrence sur n que

$$u\left(\sum_{j=1}^{n} t_{1j}x_{1j}, \ldots, \sum_{j=1}^{n} t_{rj}x_{rj}\right) = \sum_{j_1 \cdots j_r} t_{1j_1} \cdots t_{rj_r}\, u(x_{1j_1}, \ldots, x_{rj_r})$$

où les t_{ij} sont réels et les x_{ij} sont dans E_i, $i = 1, \ldots, r$, la somme du second membre étant étendue à tous les systèmes $(j_1, \ldots, j_r)$ avec $1 \leqslant j_i \leqslant n$.

[2] Dans l'exercice 1 prenons $E_1 = \cdots = E_r = E$ et $F = \mathbb{R}$; supposons que $\dim E = N \geqslant 1$ et $\{e_1, \ldots, e_N\}$ est une base de E. Si $u \in \mathcal{A}_N(E)$ vérifier que

$$u = (x_1, \ldots, x_N) = c \det[t_{ij}]$$

où $c = u(e_1, \ldots, e_N)$ et $x_i = \sum_{j=1}^{N} t_{ij}e_j$, $i = 1, \ldots, N$.

[3] Soient E un espace vectoriel réel de dimension finie $N \geqslant 1$ et $T : E \to E$ une application linéaire. Démontrer qu'il existe un unique nombre réel $\delta = \delta(T)$ tel que pour tout $u \in \mathcal{A}_N(E)$, $u \neq 0$, on a

$$u(Tx_1, \ldots, Tx_N) = \delta u(x_1, \ldots, x_N)$$

pour tout $(x_1, \ldots, x_N) \in E^N$. Si T correspond à la matrice $[t_{ij}]$ par rapport à la base $\{e_1, \ldots, e_N\}$ de E démontrer que $\delta = \det[t_{ij}]$; en déduire que $\det[t_{ij}]$ est indépendant du choix de la base $\{e_1, \ldots, e_N\}$ justifiant l'écriture $\det(T) = \det[t_{ij}]$.

[4] Déduire de l'exercice 3 que si T_1, T_2 sont deux applications linéaires de E dans E (E étant comme dans l'exercice 3) alors

$$\det(T_2T_1) = \det(T_1T_2) = \det(T_1)\det(T_2)\,.$$

On en conclut que pour tout $T : E \to E$ linéaire et bijective $\det(T) \neq 0$; démontrer que si T n'est pas bijective, $\det T = 0$ en raisonnant comme suit : alors il existe une base $\{e_1, \ldots, e_N\}$ de E telle que $Te_1 = 0$; si $u \in \mathcal{A}_N(E)$, $u \neq 0$, calculer $u(Te_1, \ldots, Te_N)$.

[5] Soient $\alpha_1, \ldots, \alpha_k$, k formes linéaires indépendantes sur $\mathbb{R}^N$. Démontrer qu'une forme linéaire α sur $\mathbb{R}^N$ est linéairement dépendante des α_i si et seulement si

$$\alpha \wedge \alpha_1 \wedge \cdots \wedge \alpha_k = 0\,.$$

7.6 FORMES DIFFÉRENTIELLES

7.6.1 Introduction

Dans cette section nous définissons les formes différentielles dans $\mathbb{R}^N$ précisément et nous démontrons les propriétés de ces formes énumérées formellement dans la section 7.1.

7.6.2 Définition des formes différentielles

Soit Ω une partie ouverte non vide de $\mathbb{R}^N$. Si p est un entier $\geqslant 0$, une ***forme différentielle extérieure de degré*** p dans Ω (brièvement, une ***p-forme différentielle*** ou encore une ***p-forme*** dans Ω) est une application

$$\omega : \Omega \to \mathcal{A}_p(\mathbb{R}^N)\,.$$

Si α et β sont deux p-formes dans Ω, on définit $\alpha + \beta$ et $t\alpha$ ($t \in \mathbb{R}$) par les opérations correspondantes à chaque point $a \in \Omega$, ainsi

$$(\alpha + \beta)(a) = \alpha(a) + \beta(a)\,, \quad (t\alpha)(a) = t\alpha(a)\,;$$

de même, si α est une p-forme et β une q-forme dans Ω, on définit $\alpha \wedge \beta$ une $(p+q)$-forme dans Ω par la formule

$$(\alpha \wedge \beta)(a) = \alpha(a) \wedge \beta(a)\,, \quad a \in \Omega\,.$$

Rappelons que, pour $p = 0$, $\mathcal{A}_0(\mathbb{R}^N) = \mathbb{R}$; donc, une 0-forme dans Ω est une fonction réelle définie dans Ω. Une 1-forme s'appelle aussi une ***forme de Pfaff***.

Si $f : \Omega \to \mathbb{R}$ est une fonction dérivable (on dira que c'est une 0-forme dérivable dans Ω), nous écrivons $\mathrm{d}f$ pour f' ; ainsi,

$$\mathrm{d}f : \Omega \longrightarrow \mathcal{L}(\mathbb{R}^N) = (\mathbb{R}^N)' = \mathcal{A}_1(\mathbb{R}^N) \sim \mathbb{R}^N$$

est une 1-forme dans Ω. Si

$$x_i : \Omega \longrightarrow \mathbb{R}$$

est la ***fonction coordonnée*** définie par

$$x_i(a_1, \ldots, a_N) = a_i\,, \quad i = 1, \ldots, N\,,$$

on aura les 1-formes $\mathrm{d}x_1, \ldots, \mathrm{d}x_N$ dans Ω définies par

$$\mathrm{d}x_i(a)(\xi) = \xi_i\,, \quad i = 1, \ldots, N\,,$$

où $a \in \Omega$, $\xi = (\xi_1, \ldots, \xi_N) \in \mathbb{R}^N$.

Evidemment, pout tout $a \in \Omega$, $\{\mathrm{d}x_1(a), \ldots, \mathrm{d}x_N(a)\}$ est une base de $\mathcal{A}_1(\mathbb{R}^N)$; ainsi, si ω est une p-forme différentielle dans Ω, on a d'après le paragraphe 7.5.3 que

$$\omega(a) = \sum_{1 \leqslant i_1 < \ldots < i_p \leqslant N} f_{i_1, \ldots, i_p}(a)\, \mathrm{d}x_{i_1}(a) \wedge \cdots \wedge \mathrm{d}x_{i_p}(a), \quad a \in \Omega,$$

où les $f_{i_1, \ldots, i_p}(a)$ sont uniquement déterminés. Pour abréger, nous écrivons (en supprimant l'argument $a \in \Omega$),

$$\mathrm{d}x_I = \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \quad \text{si } I = \{i_1 < \ldots < i_p\}$$

et

$$\omega = \sum_{I \in S(p;N)} f_I\, \mathrm{d}x_I \tag{7.31}$$

où $S(p;N)$ est l'ensemble des suites strictement croissantes d'entiers dans $\{1, \ldots, N\}$ de longueur p et les f_I sont des fonctions réelles définies dans Ω. Selon la notation de la section 7.5 on a

$$\mathrm{d}x_I(a) = \mathrm{d}x_{i_1}(a) \wedge \cdots \wedge \mathrm{d}x_{i_p}(a) = u_{i_1} \wedge \cdots \wedge u_{i_p} = u^I$$

si $u_j = \mathrm{d}x_j(a)$, $j = 1, \ldots, N$, $I = \{i_1 < \ldots < i_p\} \in S(p;N)$. Nous dirons que (7.31) est l'***écriture canonique*** de la p-forme ω et que les f_I sont les ***composantes*** de ω. La p-forme ω est de ***classe*** C^r dans Ω si toutes ses composantes sont de classe C^r dans Ω ; le ***support*** de ω (noté $\operatorname{supp}\omega$) est défini comme suit : soit

$$A = \{x : f_I(x) \neq 0 \text{ pour un } I \in S(p;N)\} = \bigcup_{I \in S(p;N)} \{x : f_I(x) \neq 0\} ;$$

alors $\operatorname{supp}\omega = \overline{A}$ où ω possède l'écriture canonique (7.31). Dire que $\operatorname{supp}\omega$ est compact revient à dire qu'il existe une partie compacte K telle que $f_I(x) = 0$ si $x \in K^c$ pour tout $I \in S(p;N)$.

7.6.3 Produit extérieur des formes différentielles

Soit Ω une partie ouverte non vide de $\mathbb{R}^N$. Soient α une p-forme et β une q-forme dans Ω ayant les écritures canoniques

$$\alpha = \sum_{I \in S(p;N)} a_I\, \mathrm{d}x_I, \quad \beta = \sum_{J \in S(q;N)} b_J\, \mathrm{d}x_J .$$

D'après (7.28) on a

$$\alpha \wedge \beta = \sum_{I,J} a_I b_J \, \mathrm{d}x_I \wedge \mathrm{d}x_J$$

où

$$\mathrm{d}x_I \wedge \mathrm{d}x_J = \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \wedge \mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q}$$

si

$$I = \{i_1 < \ldots < i_p\}, \quad J = \{j_1 < \ldots < j_q\}.$$

Il est clair que le produit extérieur des formes différentielles dans Ω possède les propriétés algébriques du produit extérieur des formes multilinéaires alternées présentées dans la section 7.5. Ainsi, les formules concernant le produit extérieur des formes différentielles données dans le paragraphe 7.1.2 sont des conséquences de la proposition 7.5.4.

7.6.4 Différentiation extérieure

Soit Ω une partie ouverte non vide de $\mathbb{R}^N$. Si f est une 0-forme de classe C^r $(r \geqslant 1)$ dans Ω, nous définissons

$$\mathrm{d}f = \mathrm{D}_1 f \, \mathrm{d}x_1 + \cdots + \mathrm{D}_N f \, \mathrm{d}x_N$$

comme une 1-forme (de classe C^{r-1}) dans Ω. Si ω est une p-forme de classe C^r, $r \geqslant 1$ dans Ω, ayant l'écriture canonique

$$\omega = \sum_{I \in S(p;N)} f_I \, \mathrm{d}x_I$$

nous définissons

$$\begin{aligned} \mathrm{d}\omega &= \sum_{I \in S(p;N)} \mathrm{d}f_I \wedge \mathrm{d}x_I \\ &= \sum_{\{i_1 < \ldots < i_p\} \in S(p;N)} \mathrm{d}f_{i_1 \ldots i_p} \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \ ; \end{aligned} \tag{7.32}$$

$\mathrm{d}\omega$ est une $(p+1)$-forme (de classe C^{r-1}) dans Ω. Dans le paragraphe 7.1.3 nous avons déjà pratiqué les règles du calcul de $\mathrm{d}\omega$. Ici nous donnons les démonstrations de la validité de ces règles, en particulier pour celles contenues dans (7.2) et (7.3).

La linéarité de l'opération d est évidente si α et β sont des p-formes (de classe C^1) dans Ω,

$$\mathrm{d}(\alpha + \beta) = \mathrm{d}\alpha + \mathrm{d}\beta \ ; \quad \mathrm{d}(t\alpha) = t(\mathrm{d}\alpha), \ t \in \mathbb{R}.$$

Démontrons que

$$\mathrm{d}(\alpha \wedge \beta) = (\mathrm{d}\alpha) \wedge \beta + (-1)^p \alpha \wedge \mathrm{d}\beta$$

où α est une p-forme et β est une q-forme (de classe C^1) dans Ω. Si $p = q = 0$, alors α et β sont deux fonctions de classe C^1 et $\alpha \wedge \beta = \alpha\beta$; dans ce cas, la formule ci-dessus est simplement celle de la dérivation du produit de deux fonctions ; le cas général en est une conséquence. En effet, si

$$\alpha = \sum_{I \in S(p;N)} a_I \, \mathrm{d}x_I \,, \quad \beta = \sum_{J \in S(q;N)} b_J \, \mathrm{d}x_J$$

on a

$$\alpha \wedge \beta = \sum_{I,J} a_I b_J \, \mathrm{d}x_I \wedge \mathrm{d}x_J$$

d'où

$$\begin{aligned}
\mathrm{d}(\alpha \wedge \beta) &= \sum_{I,J} (a_I \, \mathrm{d}b_J + b_J \, \mathrm{d}a_I) \wedge \mathrm{d}x_I \wedge \mathrm{d}x_J \\
&= \sum_{I,J} \{ b_J \, \mathrm{d}a_I \wedge \mathrm{d}x_I \wedge \mathrm{d}x_J + a_I \, \mathrm{d}b_J \wedge \mathrm{d}x_I \wedge \mathrm{d}x_J \} \\
&= (\mathrm{d}\alpha \wedge \beta) + \sum_{I,J} a_I (-1)^p \, \mathrm{d}x_I \wedge \mathrm{d}b_J \wedge \mathrm{d}x_J \\
&= (\mathrm{d}\alpha \wedge \beta) + (-1)^p (\alpha \wedge \mathrm{d}\beta)
\end{aligned}$$

car

$$\mathrm{d}b_J \wedge \mathrm{d}x_I = (-1)^{1 \cdot p} \, \mathrm{d}x_I \wedge \mathrm{d}b_J = (-1)^p \, \mathrm{d}x_I \wedge \mathrm{d}b_J$$

$$\alpha \wedge \mathrm{d}\beta = \sum_{I,J} a_I \, \mathrm{d}x_I \wedge \mathrm{d}b_J \wedge \mathrm{d}x_J \,.$$

Démontrons maintenant la propriété fondamentale :

$$\mathrm{d}(\mathrm{d}\alpha) = 0$$

si α est une p-forme de classe C^2 dans Ω. Si $p = 0$, $\alpha : \Omega \to \mathbb{R}$ est une fonction de classe C^2 ; donc,

$$\mathrm{d}\alpha = \sum_{i=1}^{N} \mathrm{D}_i \alpha \, \mathrm{d}x_i$$

et

$$\mathrm{d}(\mathrm{d}\alpha) = \sum_{i=1}^{N} \mathrm{d}(\mathrm{D}_i\alpha) \wedge \mathrm{d}x_i = \sum_{i=1}^{N} \left(\sum_{j=1}^{N} \mathrm{D}_j\mathrm{D}_i\alpha \, \mathrm{d}x_j \right) \wedge \mathrm{d}x_i$$

$$= \sum_{1 \leqslant i,\, j \leqslant N} (\mathrm{D}_j\mathrm{D}_i\alpha) \, \mathrm{d}x_j \wedge \mathrm{d}x_i$$

$$= \sum_{1 \leqslant i < j \leqslant N} \{\mathrm{D}_i\mathrm{D}_j\alpha - \mathrm{D}_j\mathrm{D}_i\alpha\} \, \mathrm{d}x_i \wedge \mathrm{d}x_j = 0$$

car $\mathrm{d}x_i \wedge \mathrm{d}x_i = 0$, $\mathrm{d}x_j \wedge \mathrm{d}x_i = -\mathrm{d}x_i \wedge \mathrm{d}x_j$ et $\mathrm{D}_j\mathrm{D}_i\alpha = \mathrm{D}_i\mathrm{D}_j\alpha$. Pour $p \geqslant 1$, soit

$$\alpha = \sum_{I \in S(p;N)} a_I \, \mathrm{d}x_I \ ;$$

alors

$$\mathrm{d}\alpha = \sum_{I} \mathrm{d}a_I \wedge \mathrm{d}x_I \ ;$$

notons que

$$\mathrm{d}(\mathrm{d}a_I) = 0$$

et

$$\mathrm{d}(\mathrm{d}x_I) = \mathrm{d}(1 \cdot \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}), \quad I = \{i_1 < \ldots < i_p\}$$
$$= \mathrm{d}1 \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} = 0 .$$

Donc

$$\mathrm{d}(\mathrm{d}\alpha) = \sum_{I} \mathrm{d}(\mathrm{d}a_I \wedge \mathrm{d}x_I) = \sum_{I} \{\mathrm{d}(\mathrm{d}a_I) \wedge \mathrm{d}x_I + (-1)\, \mathrm{d}a_I \wedge \mathrm{d}(\mathrm{d}x_I)\} = 0 .$$

7.6.5 Changement de variables

Nous reprenons les notations du paragraphe 7.1.5 ; soit

$$\varphi : U \longrightarrow V$$

une fonction de classe C^r, $r \geqslant 1$, où $U \subset \mathbb{R}^N$, $V \subset \mathbb{R}^M$, U et V ouverts et $\varphi = (\varphi_1, \ldots, \varphi_M)$; les fonctions coordonnées de $\mathbb{R}^N$ sont notées $x_1, \ldots, x_N$, celles de $\mathbb{R}^M$ sont $y_1, \ldots, y_M$. Si

$$\omega = \sum_{I \in S(p;M)} f_I \, \mathrm{d}y_I$$

est une p-forme dans V, on obtient $\varphi^*\omega$, une p-forme dans U, par la formule

$$\varphi^*\omega = \sum_{I \in S(p;M)} (f_I \circ \varphi) \, \mathrm{d}\varphi_I$$

où

$$\mathrm{d}\varphi_I = \mathrm{d}\varphi_{i_1} \wedge \cdots \wedge \mathrm{d}\varphi_{i_p}$$

si $I = \{i_1 < \ldots < i_p\}$, $1 \leqslant i_1 < \ldots < i_p \leqslant M$ et

$$\mathrm{d}\varphi_i = \sum_{j=1}^{N} (\mathrm{D}_j \varphi_i)\, \mathrm{d}x_j \,.$$

(i) *φ^* commute avec les opérations algébriques sur les formes dans V.*

En effet, la linéarité de $\omega \mapsto \varphi^*\omega$ est évidente. Si

$$\alpha = \sum_I a_I \,\mathrm{d}y_I \,, \quad \beta = \sum_J b_J \,\mathrm{d}y_J$$

sont deux formes dans V, α une p-forme, β une q-forme alors

$$\alpha \wedge \beta = \sum_{I,J} a_I b_J \,\mathrm{d}y_I \wedge \mathrm{d}y_J$$

$$\varphi^*(\alpha \wedge \beta) = \sum_{I,J} (a_I b_J) \circ \varphi \,\mathrm{d}\varphi_I \wedge \mathrm{d}\varphi_J$$

$$= \sum_{I,J} (a_I \circ \varphi)(b_J \circ \varphi)\, \mathrm{d}\varphi_I \wedge \mathrm{d}\varphi_J = \varphi^*(\alpha) \wedge \varphi^*(\beta)$$

où $I \in S(p; M)$, $J \in S(q; M)$.

(ii) *φ^* commute avec* d *: on a*

$$\mathrm{d}(\varphi^*\omega) = \varphi^*(\mathrm{d}\omega)$$

si φ est de classe C^2 et ω est une p-forme de classe C^1 dans V.

En effet, on vérifie ceci d'abord pour $p = 0$ auquel cas $\omega = f : V \to \mathbb{R}$, une fonction de classe C^1 ; alors $\varphi^* f = f \circ \varphi$ et

$$\mathrm{d}(\varphi^* f) = \mathrm{D}_1(f \circ \varphi)\, \mathrm{d}x_1 + \cdots + \mathrm{D}_N(f \circ \varphi)\, \mathrm{d}x_N$$

$$= \sum_{i=1}^{N} \left\{ \sum_{j=1}^{M} (\mathrm{D}_j f) \circ \varphi \cdot \mathrm{D}_i \varphi_j \right\} \mathrm{d}x_i = \sum_{j=1}^{M} (\mathrm{D}_j f) \circ \varphi \sum_{i=1}^{N} \mathrm{D}_i \varphi_j \,\mathrm{d}x_i$$

$$= \sum_{j=1}^{M} (\mathrm{D}_j f) \circ \varphi \,\mathrm{d}\varphi_j = \varphi^* \left(\sum_{j=1}^{M} \mathrm{D}_j f \,\mathrm{d}y_j \right) = \varphi^*(\mathrm{d}f)$$

où l'on a utilisé la formule de dérivation des fonctions composées pour avoir

$$\mathrm{D}_i(f \circ \varphi) = (\mathrm{D}_1 f) \circ \varphi \cdot \mathrm{D}_i \varphi_1 + \cdots + (\mathrm{D}_M f) \circ \varphi \cdot \mathrm{D}_i \varphi_M \,, \quad 1 \leqslant i \leqslant N \,,$$

car $(f \circ \varphi)(x) = f\big(\varphi_1(x), \ldots, \varphi_M(x)\big)$, $x \in U$. Dans cette étape il suffisait que φ soit de classe C^1.

Si ω est une p-forme dans V du type (monôme) $\omega = f\,\mathrm{d}y_{i_1} \wedge \cdots \wedge \mathrm{d}y_{i_p}$, $f : V \to \mathbb{R}$ de classe C^1, on a

$$\varphi^*\omega = (f \circ \varphi)\,\mathrm{d}\varphi_{i_1} \wedge \cdots \wedge \mathrm{d}\varphi_{i_p} ;$$

notons que $\varphi^*\omega$ est de classe C^1 car φ est supposé être de classe C^2. Donc,

$$\mathrm{d}(\varphi^*\omega) = \mathrm{d}(f \circ \varphi) \wedge \mathrm{d}\varphi_{i_1} \wedge \cdots \wedge \mathrm{d}\varphi_{i_p}$$

d'après la règle de calcul de $\mathrm{d}(\alpha \wedge \beta)$ et le fait que $\mathrm{d}(\mathrm{d}\varphi_i) = 0$ si φ_i est de classe C^2 ; ainsi,

$$\begin{aligned}\mathrm{d}(\varphi^*\omega) &= \mathrm{d}(\varphi^* f) \wedge \mathrm{d}\varphi_{i_1} \wedge \cdots \wedge \mathrm{d}\varphi_{i_p} = \varphi^*(\mathrm{d}f) \wedge \mathrm{d}\varphi_{i_1} \wedge \cdots \wedge \mathrm{d}\varphi_{i_p} \\ &= \varphi^*\left(\sum_{j=1}^{M} \mathrm{D}_j f\,\mathrm{d}y_j\right) \wedge \mathrm{d}\varphi_{i_1} \wedge \cdots \wedge \mathrm{d}\varphi_{i_p} \\ &= \sum_{j=1}^{M} (\mathrm{D}_j f) \circ \varphi\,\mathrm{d}\varphi_j \wedge \mathrm{d}\varphi_{i_1} \wedge \cdots \wedge \mathrm{d}\varphi_{i_p} = \varphi^*(\mathrm{d}\omega)\end{aligned}$$

car

$$\mathrm{d}\omega = \mathrm{d}f \wedge \mathrm{d}y_{i_1} \wedge \cdots \wedge \mathrm{d}y_{i_p} = \sum_{j=1}^{M} \mathrm{D}_j f\,\mathrm{d}y_j \wedge \mathrm{d}y_{i_1} \wedge \cdots \wedge \mathrm{d}y_{i_p}\,.$$

Par linéarité, on obtient la formule $\mathrm{d}(\varphi^*\omega) = \varphi^*(\mathrm{d}\omega)$ pour toute p-forme de classe C^1 dans V.

(iii) *Transitivité des changements de variables : si*

$$U \xrightarrow{\varphi} V \xrightarrow{\psi} W$$

avec $U \subset \mathbb{R}^N$, $V \subset \mathbb{R}^M$, $W \subset \mathbb{R}^P$ trois ensembles ouverts, φ et ψ de classe C^1, alors

$$(\psi \circ \varphi)^*\omega = \varphi^*(\psi^*\omega)$$

pour toute p-forme ω dans W.

En effet, en écrivant les fonctions coordonnées dans W comme $z_1, \ldots, z_P$, il suffit de vérifier la formule pour

$$\omega = f\,\mathrm{d}z_{i_1} \wedge \cdots \wedge \mathrm{d}z_{i_p}$$

où $f : W \to \mathbb{R}$. Posons $\psi \circ \varphi = \gamma$, $\gamma = (\gamma_1, \ldots, \gamma_P)$, $\varphi = (\varphi_1, \ldots, \varphi_M)$, $\psi = (\psi_1, \ldots, \psi_P)$; alors

$$\gamma_i = \psi_i \circ \varphi = \varphi^*\psi_i\,, \quad \mathrm{d}\gamma_i = \mathrm{d}(\varphi^*\psi_i) = \varphi^*(\mathrm{d}\psi_i)$$

et

$$\begin{aligned}(\psi\circ\varphi)^*(\omega) &= \gamma^*(f\,\mathrm{d}z_{i_1}\wedge\cdots\wedge\mathrm{d}z_{i_p}) = (f\circ\gamma)\,\mathrm{d}\gamma_{i_1}\wedge\cdots\wedge\mathrm{d}\gamma_{i_p}\\ &= (f\circ\gamma)\varphi^*(\mathrm{d}\psi_{i_1})\wedge\cdots\wedge\varphi^*(\mathrm{d}\psi_{i_p})\\ &= (f\circ\psi\circ\varphi)\varphi^*(\mathrm{d}\psi_{i_1}\wedge\cdots\wedge\mathrm{d}\psi_{i_p})\\ &= \varphi^*\{(f\circ\psi)\,\mathrm{d}\psi_{i_1}\wedge\cdots\wedge\mathrm{d}\psi_{i_P}\}\\ &= \varphi^*\{\psi^*(f\,\mathrm{d}z_{i_1}\wedge\cdots\wedge\mathrm{d}z_{i_p})\} = \varphi^*(\psi^*\omega)\,;\end{aligned}$$

ceci établit la formule donnée.

Pour conclure ce paragraphe, écrivons $\varphi^*\omega$ dans son écriture canonique. Soit

$$\omega = \sum_{I\in S(p;M)} f_I\,\mathrm{d}y_I$$

une p-forme dans $V\subset\mathbb{R}^M$; alors

$$\varphi^*\omega = \sum_{I\in S(p;M)} (f_I\circ\varphi)\,\mathrm{d}\varphi_I$$

avec $\mathrm{d}\varphi_I = \mathrm{d}\varphi_{i_1}\wedge\cdots\wedge\mathrm{d}\varphi_{i_p}$, $I=\{i_1<\ldots<i_p\}$. Puisque

$$\mathrm{d}\varphi_i = \sum_{j=1}^{N}(\mathrm{D}_j\varphi_i)\,\mathrm{d}x_j\,,\quad 1\leqslant i\leqslant M\,,$$

on a pour $p\leqslant N$, d'après (7.30), que

$$\mathrm{d}\varphi_{i_1}\wedge\cdots\wedge\mathrm{d}\varphi_{i_p} = \sum_{1\leqslant j_1<\ldots<j_p\leqslant N}\frac{\partial(\varphi_{i_1},\ldots,\varphi_{i_p})}{\partial(x_{j_1},\ldots,x_{j_p})}\,\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_p}\,.$$

Donc

$$\varphi^*\omega = \sum_{\substack{I\in S(p;M)\\ J\in S(p;N)}}(f_I\circ\varphi)\frac{\partial(\varphi_{i_1},\ldots,\varphi_{i_p})}{\partial(x_{j_1},\ldots,x_{j_p})}\,\mathrm{d}x_J = \sum_{J\in S(p;N)} g_J\,\mathrm{d}x_J \qquad (7.33)$$

où

$$g_J = \sum_{I\in S(p;M)}(f_I\circ\varphi)\frac{\partial(\varphi_{i_1},\ldots,\varphi_{i_p})}{\partial(x_{j_1},\ldots,x_{j_p})}$$

avec $I=\{i_1<\ldots<i_p\}\in S(p;M)$, $J=\{j_1<\cdots<j_p\}\in S(p;N)$.

Si $p>N$, $\varphi^*\omega=0$ ce qui est banal car toute p-forme dans $U\subset\mathbb{R}^N$ est nulle si $p>N$.

7.7 ORIENTATION

7.7.1 Variété différentielle orientée

Soit X une sous-variété différentielle de dimension $k \geqslant 1$ dans $\mathbb{R}^N$ (de classe C^1). Soient (ψ_1, U_1), (ψ_2, U_2) deux paramétrages locaux de deux voisinages V_1, V_2 d'un même point $a \in X$ (où V_1, V_2 sont des ensembles X-ouverts contenant a ; cf. § 7.3.1) ; nous dirons que ces deux paramétrages ont la ***même orientation*** (ou donnent la même orientation) si le changement de paramétrage de ψ_1 à ψ_2 donné par $\varphi = \psi_2^{-1} \circ \psi_1$ est direct (cf. § 7.3.2), c'est-à-dire si $\det \varphi'(u) > 0$ pour tout u dans $W_1 = \psi_1^{-1}(V_1 \cap V_2)$; nous dirons que ces deux paramétrages ont les ***orientations opposées*** (ou bien, l'orientation d'un paramétrage est opposée à celle de l'autre) si le changement de paramétrage φ est opposé, c'est-à-dire si $\det \varphi'(u) < 0$ pour $u \in W_1$.

Rappelons du paragraphe 7.3.1 qu'un atlas pour X est une famille $\mathcal{A}$ de paramétrages locaux $(\psi_i, U_i)_{i \in I}$ de X tels que

$$\bigcup_{i \in I} V_i = X$$

où $\psi_i(U_i) = V_i$; on dit que $\mathcal{A}$ est ***orienté*** si pour chaque couple i, j de I ou bien $V_i \cap V_j = \emptyset$ ou bien $V_i \cap V_j \neq \emptyset$ et les deux paramétrages (ψ_i, U_i), (ψ_j, U_j) donnent la même orientation. On dit que X est ***orientable*** si X possède un atlas orienté ; X est dit ***orienté*** par le choix d'un tel atlas.

Soit X orienté par un atlas $\mathcal{A}$; un paramétrage local (ψ, U) quelconque est dit ***positivement orienté*** (sous-entendu, par rapport à l'orientation de X fixée par $\mathcal{A}$) si pour tout choix d'un paramétrage local (ψ_i, U_i) de $\mathcal{A}$, les deux paramétrages (ψ, U), (ψ_i, U_i) donnent la même orientation dès que

$$\psi(U) \cap \psi_i(U_i) \neq \emptyset .$$

On dit que deux atlas orientés $\mathcal{A}$ et $\widetilde{\mathcal{A}}$ pour X donnent la ***même orientation*** à X (ou que $\mathcal{A}$ et $\widetilde{\mathcal{A}}$ sont ***équivalents***) si pour chaque choix de $(\psi, U) \in \mathcal{A}$, $(\widetilde{\psi}, \widetilde{U}) \in \widetilde{\mathcal{A}}$ avec $\psi(U) \cap \widetilde{\psi}(\widetilde{U}) \neq \emptyset$, on a que (ψ, U), $(\widetilde{\psi}, \widetilde{U})$ donnent la même orientation ; $\mathcal{A}$ et $\widetilde{\mathcal{A}}$ donnent les ***orientations opposées*** à X (ou bien, $\mathcal{A}$ et $\widetilde{\mathcal{A}}$ sont ***opposés***) si, sous les mêmes conditions, (ψ, U), $(\widetilde{\psi}, \widetilde{U})$ ont les orientations opposées.

Soit X orienté par un atlas $\mathcal{A}$; *il existe un atlas* $\widetilde{\mathcal{A}}$ *pour* X *opposé à* $\mathcal{A}$. En effet, soit $r : \mathbb{R}^k \to \mathbb{R}^k$ défini par

$$r(u_1, u_2, \ldots, u_k) = (-u_1, u_2, \ldots, u_k) ;$$

nous dirons que r ***renverse*** les orientations. Notons que $r \circ r = \mathrm{id}$ et que $\det r'(u) = -1$ pour $u \in \mathbb{R}^k$. Soit $\mathcal{A}$ l'atlas orienté formé de (ψ_i, U_i), $i \in I$; posons

$$\widetilde{U}_i = r(U_i) , \quad \widetilde{\psi}_i = \psi_i \circ r ;$$

puisque $\widetilde{\psi}_i(\widetilde{U}_i) = \psi_i(U_i)$, on vérifie immédiatement que $\widetilde{\mathcal{A}} = \{(\widetilde{\psi}_i, \widetilde{U}_i)\}_{i\in I}$ est un atlas ; cet atlas est orienté car

$$\widetilde{\psi}_j^{-1} \circ \widetilde{\psi}_i = r \circ (\psi_j^{-1} \circ \psi_i) \circ r\,,$$

d'où, en utilisant $\big(\det r'\big)^2 = 1$, on a que

$$\det\big(\widetilde{\psi}_j^{-1} \circ \widetilde{\psi}_i\big)' = \det\big(\psi_j^{-1} \circ \psi_i\big)' \circ r > 0$$

(dès que $\psi_i(U_i) \cap \psi_j(U_j) \neq \emptyset$). De plus, si $\psi_i(U_i) \cap \psi_j(U_j) \neq \emptyset$, on aura

$$\psi_j^{-1} \circ \widetilde{\psi}_i = \psi_j^{-1} \circ \psi_i \circ r$$

avec

$$\det\big(\psi_j^{-1} \circ \widetilde{\psi}_i\big)' = -\det\big(\psi_j^{-1} \circ \psi_i\big)' \circ r < 0$$

ce qui prouve que $\mathcal{A}$ et $\widetilde{\mathcal{A}}$ sont opposés.

On verra dans la suite que la définition de l'orientation donnée ici est équivalente à la définition de l'orientation donnée pour une surface dans $\mathbb{R}^3$ en utilisant l'existence d'une normale unitaire continue. Ainsi, nous savons qu'il existe des surfaces dans $\mathbb{R}^3$ (comme le ruban de Moebius) qui ne sont pas orientables.

Les exemples les plus simples des variétés différentielles orientables sont les morceaux ; ces variétés sont données par un seul paramétrage $\psi : U \to \mathbb{R}^N$ avec $X = \psi(U)$, U étant un domaine dans $\mathbb{R}^k$. Il est clair aussi que si X est une variété différentielle orientable, toute partie non vide ouverte dans X est aussi une variété différentielle orientable. Une variété différentielle de dimension 1 dans $\mathbb{R}^N$ (ce qui est une réunion des courbes régulières deux à deux disjointes – mais ceci exige une démonstration) est toujours orientable ; elle possède exactement deux orientations, une opposée à l'autre, si elle est connexe. Nous ne donnons pas une preuve de la première affirmation ; elle est une conséquence du fait que chaque variété différentielle connexe de classe C^1 de dimension 1 dans $\mathbb{R}^N$ est C^1-difféomorphe à une courbe régulière dans $\mathbb{R}^N$; en fait, elle est C^1-difféomorphe à un cercle si la variété est compacte et est C^1-difféomorphe à un intervalle ouvert de $\mathbb{R}$ si la variété est non compacte ; voir Milnor [26, pp. 55-57] pour une démonstration (simple) de ces affirmations qui semblent être intuitivement évidentes. On démontrera dans la suite que toute variété différentielle connexe *orientable* possède exactement deux orientations, une opposée à l'autre.

7.7.2 Orientation des espaces vectoriels

Soit E un espace vectoriel réel de dimension $n \geqslant 1$. Si $a = \{a_1, \ldots, a_n\}$ et $b = \{b_1, \ldots, b_n\}$ sont deux bases ordonnées de E, on dira que a est ***équivalente*** à b ($a \sim b$) si $\det[t_{ij}] > 0$ où

$$b_i = \sum_{j=1}^{n} t_{ij} a_j\,, \quad i = 1, \ldots, n\,;$$

si $\det[t_{ij}] < 0$, on dira que a et b sont ***opposées*** ($a \not\sim b$). Il est clair que $a \sim a$; si $a \sim b$ on a aussi $b \sim a$ car

$$a_i = \sum_{j=1}^{n} s_{ij} b_j\,, \quad i = 1, \ldots, n\,,$$

avec $[s_{ij}] = [t_{ij}]^{-1}$ d'où

$$\det[s_{ij}] > 0 \quad \text{si} \quad \det[t_{ij}] > 0\,;$$

de plus, si $a \sim b$ et $b \sim c = \{c_1, \ldots, c_n\}$, une autre base ordonnée de E alors $a \sim c$ ce que l'on démontre facilement en utilisant les propriétés de multiplication des déterminants. Ainsi, $\sim$ est une relation d'équivalence dans l'ensemble des bases ordonnées de E. Fixons une base ordonnée a de E et posons

$$\mathcal{B} = \{b : b \sim a\}\,, \quad \mathcal{B}' = \{b : b \not\sim a\}$$

où b est une base ordonnée quelconque de E ; il est clair que toute base ordonnée de E est soit dans $\mathcal{B}$, soit dans $\mathcal{B}'$, $\mathcal{B}$ étant la classe d'équivalence des bases ordonnées contenant a. Par ailleurs, on démontre facilement que deux bases ordonnées de E sont équivalentes si et seulement si les deux sont dans $\mathcal{B}$ ou les deux sont dans $\mathcal{B}'$; deux bases ordonnées de E sont opposées si et seulement si une est dans $\mathcal{B}$ et l'autre est dans $\mathcal{B}'$.

On dit que E est ***orienté*** si l'on a distingué un des deux ensembles $\mathcal{B}$, $\mathcal{B}'$ comme celui qui contient les bases ordonnées dites ***positives***, l'autre ensemble alors étant formé des bases ordonnées ***négatives***. Ainsi, E possède deux ***orientations***, l'une correspondant au choix de $\mathcal{B}$ comme l'ensemble des bases ordonnées positives, l'autre correspondant au choix du $\mathcal{B}'$ comme l'ensemble des bases ordonnées négatives ; on écrit $\mathcal{O}_E = \mathcal{B}$ dans le premier cas et $\mathcal{O}_E = \mathcal{B}'$ dans le second et on dit que l'orientation de E est donnée par $\mathcal{O}_E$ ou encore que l'orientation de E est $\mathcal{O}_E$.

Si $n \geqslant 2$ et $a = \{a_1, a_2, \ldots, a_n\}$ est une base ordonnée de E alors $a' = \{a_2, a_1, \ldots, a_n\}$ est une base ordonnée opposée à a et

$$\mathcal{B} = \{b : b \sim a\}\,, \quad \mathcal{B}' = \{b : b \sim a'\}\,.$$

Si $\dim E = 1$, $\mathcal{B}$ est déterminé par le choix d'un vecteur $a \neq 0$, $\mathcal{B}'$ étant alors déterminé par $a' = -a$.

Pour l'espace vectoriel $\mathbb{R}^N$, l'orientation ***canonique*** (ou ***usuelle***) est déterminée en choisissant la base ordonnée canonique $e = \{e_1, \ldots, e_N\}$ comme étant positive où $e_1 = (1, 0, \ldots, 0)$, $\ldots$, $e_N = (0, 0, \ldots, 1)$; on a alors

$$\mathcal{B} = \{b : b \sim e\}\,, \quad \mathcal{B}' = \{b : b \not\sim e\}\,.$$

Pour $N = 1$, l'orientation canonique correspond aussi à celle que l'on indique traditionnellement par des flèches sur une droite. Sans mention du contraire,

$\mathbb{R}^N$ sera orienté canoniquement ; ainsi, une base ordonnée $\{b_1, \ldots, b_N\}$ dans $\mathbb{R}^N$ est ***positive*** si et seulement si

$$\det[b_1, \ldots, b_N] > 0\,.$$

Notons que, pour orienter un espace vectoriel réel E de dimension finie, il suffit de fixer *une* base ordonnée a comme positive ; alors E est orienté en posant

$$\mathcal{O}_E = \{b : b \sim a\}$$

où b parcourt toutes les bases ordonnées de E.

Soient E et F deux espaces vectoriels réels orientés de même dimension finie $n \geqslant 1$. Un isomorphisme $u : E \to F$ (ainsi, u est linéaire et bijective) s'appelle ***positif*** si une base positive $\{b_1, \ldots, b_n\}$ de E se transforme en une base positive $\{u(b_1), \ldots, u(b_n)\}$ de F ; il est facile de démontrer qu'alors toute base positive (resp. négative) de E se transforme en une base positive (resp. négative) de F. L'isomorphisme $u : E \to F$ s'appelle ***négatif*** si une base positive de E (et alors toute base positive de E) se transforme en une base négative de F

Soient E, F et G trois espaces vectoriels réels orientés de même dimension finie. Si

$$u : E \longrightarrow F\,, \quad v : F \longrightarrow G\,, \quad w : E \longrightarrow G$$

sont trois isomorphismes tels que $w = vu$ alors la positivité de deux quelconques de ces isomorphismes entraîne la positivité du troisième. La démonstration de cette affirmation est laissée en exercice.

7.7.3 Orientation des espaces tangents

Soit X une variété différentielle de dimension $k \geqslant 1$ dans $\mathbb{R}^N$ (de classe C^1). Pour $x \in X$, soit $T_x = T_x(X)$ l'espace tangent à X en x ; T_x est un sous-espace vectoriel de $\mathbb{R}^N$ de dimension k (cf. § 7.3.3).

Supposons que $\mathcal{O}_{T_x}$ est une orientation de T_x, $x \in X$; posons $\mathcal{O} = \{\mathcal{O}_{T_x}\}_{x \in X}$. Nous dirons que $\mathcal{O}$ est une famille ***cohérente*** d'orientations en $a \in X$ s'il existe un voisinage X-ouvert V de a et un paramétrage local (ψ, U), $\psi(U) = V$, tel que, pour tout $u \in U$, l'isomorphisme

$$\psi'(u) : \mathbb{R}^k \longrightarrow T_{\psi(u)}$$

est positif, c'est-à-dire

$$\{\mathrm{D}_1\psi(u), \ldots, \mathrm{D}_k\psi(u)\} \in \mathcal{O}_{T_{\psi(u)}}\,, \quad u \in U\,;$$

on dira simplement que $\psi'(u)$ est positif pour $u \in U$. Nous dirons que $\mathcal{O}$ est ***cohérente*** si elle est cohérente partout dans X.

Démontrons que X *possède une famille cohérente d'orientations si et seulement si* X *est orientable*. En effet, soit $\mathcal{A}$ un atlas orienté pour X avec

$$\mathcal{A} = \{\psi_i, U_i\}_{i \in I}\,, \quad V_i = \psi_i(U_i)\,.$$

Si $x \in V_i$ avec $\psi_i(u) = x$, $u \in U_i$, nous fixons $\mathcal{O}_{T_x}$ en demandant que $\psi_i'(u)$ soit positif ; démontrons que ceci détermine $\mathcal{O}_{T_x}$ de manière unique. Si $x \in V_j$ avec $\psi_j(v) = x$, $v \in U_j$, alors, en posant $\varphi = \psi_j^{-1} \circ \psi_i$ (sur un domaine approprié, cf. § 7.3.2), on aura $\varphi(u) = v$ et

$$\psi_j\big(\varphi(u)\big) = \psi_i(u)\,, \quad \psi_j'(v)\,\varphi'(u) = \psi_i'(u)\,;$$

puisque $\psi_i'(u) : \mathbb{R}^k \to T_x$ est positif par définition et $\varphi'(u) : \mathbb{R}^k \to \mathbb{R}^k$ est positif car $\det \varphi'(u) > 0$, on conclut que $\psi_j'(v) : \mathbb{R}^k \to T_x$ est aussi positif (cf. § 7.7.2). Il est maintenant évident que $\mathcal{O} = \{\mathcal{O}_{T_x}\}_{x \in X}$ est une famille cohérente car pour tout $a \in X$ si $a \in V_i$, le paramétrage local (ψ_i, U_i) sera tel que $\psi_i'(u)$ est positif pour tout $u \in U_i$. D'autre part, soit $\mathcal{O} = \{\mathcal{O}_{T_x}\}_{x \in X}$ une famille cohérente d'orientations ; alors, pour tout $x \in X$, on a un voisinage X-ouvert V_x de x et un paramétrage local (ψ_x, U_x) avec $\psi_x(U_x) = V_x$ tel que $\psi_x'(u)$ est positif pour tout $x \in U_x$; on va voir que l'atlas

$$\{(\psi_x, U_x)\}_{x \in X}$$

est orienté. En effet, supposons que $V = V_x \cap V_y \neq \emptyset$; posons $\varphi = \psi_y^{-1} \circ \psi_x$, $\varphi : W_1 \to W_2$ où (fig. 7.1) :

$$W_1 = \psi_x^{-1}(V)\,, \quad W_2 = \psi_y^{-1}(V)\,;$$

on a sur W_1, les relations

$$\psi_y \circ \varphi = \psi_x\,, \quad (\psi_y' \circ \varphi)\varphi' = \psi_x'\,.$$

Puisque $\psi_x'(u)$ est positif pour tout $u \in W_1$ et $\psi_y'(u)$ est positif pour tout $u \in W_2$, $\varphi'(u)$ est positif pour tout $u \in W_1$ ce qui montre que (ψ_x, U_x), (ψ_y, V_y) ont la même orientation. ♦

Notons que l'on a démontré l'affirmation précise suivante : *chaque atlas orienté $\mathcal{A}$ pour X correspond à une unique famille d'orientation cohérente $\mathcal{O} = \{\mathcal{O}_{T_x}\}_{x \in X}$; réciproquement, $\mathcal{O}$ détermine un atlas orienté $\mathcal{A}$ pour X.*

Démontrons maintenant que si X *est une variété orientable connexe alors X possède exactement deux orientations.* En effet, de ce qui précède, on voit qu'une orientation est uniquement fixée par une famille cohérente d'orientations $\mathcal{O} = \{\mathcal{O}_{T_x}\}_{x \in X}$. Notons par $\mathcal{O}'_{T_x}$ l'orientation de T_x opposée à $\mathcal{O}_{T_x}$ et posons $\mathcal{O}' = \{\mathcal{O}'_{T_x}\}_{x \in X}$. En utilisant l'application de renversement r du paragraphe 7.7.1, on voit que $\mathcal{O}'$ est une famille cohérente d'orientations. Démontrons que si $\widetilde{\mathcal{O}} = \{\widetilde{\mathcal{O}}_{T_x}\}_{x \in X}$ est une famille cohérente d'orientations alors $\widetilde{\mathcal{O}} = \mathcal{O}$ ou $\widetilde{\mathcal{O}} = \mathcal{O}'$. A cet effet, posons

$$\varepsilon(x) = \begin{cases} 1 & \text{si } \widetilde{\mathcal{O}}_{T_x} = \mathcal{O}_{T_x} \\ -1 & \text{si } \widetilde{\mathcal{O}}_{T_x} = \mathcal{O}'_{T_x}\,. \end{cases}$$

Puisque $\widetilde{\mathcal{O}}$ est cohérente, l'application $\varepsilon : X \to \{-1, 1\}$ est localement constante ; X étant connexe, on en déduit que ε est constante (prop. 1.8.2(v)). Alors s'il existe un $a \in X$ tel que $\varepsilon(a) = 1$ on aura $\widetilde{\mathcal{O}} = \mathcal{O}$; sinon $\widetilde{\mathcal{O}} = \mathcal{O}'$. ♦

7.7.4 Orientation des hypersurfaces

Soit X une hypersurface de classe C^1 dans $\mathbb{R}^N$ (donc la dimension de X est $N-1$). Rappelons les faits suivants concernant X :

- l'espace tangent $T_x = T_x(X)$ à X en x est un sous-espace vectoriel de $\mathbb{R}^N$ de dimension $N-1$; l'espace orthogonal $T_x^\perp$ dans $\mathbb{R}^N$ étant de dimension 1 il existe exactement deux normales unitaires à X en x notées $\pm\xi(x)$; ainsi, $\|\xi(x)\| = 1$, $\xi(x) \cdot \tau = 0$ si $\tau \in T_x$;
- pour tout point $a \in X$, il existe une boule B centrée en a et une fonction $\theta : B \to \mathbb{R}$ de classe C^1 avec $\operatorname{grad}\theta(x) \neq 0$, $x \in B$, telles que

$$B \cap X = \{x \in X : \theta(x) = 0\} ;$$

posons $\widetilde{\xi}(x) = \nabla\theta(x)/\|\nabla\theta(x)\|$, $x \in B$; alors $x \mapsto \widetilde{\xi}(x)$ est continue et $\widetilde{\xi}(x)$ est une normale unitaire à X en $x \in B \cap X$.

Une fonction *continue* $\nu : X \to \mathbb{R}^N$ s'appelle un ***champ normal unitaire continu*** sur X si $\nu(x)$ est une normale unitaire à X en $x \in X$. L'objectif de ce paragraphe est d'établir que X *est orientable si et seulement s'il existe un champ normal unitaire continu sur* X.

Avant de faire la démonstration, introduisons la terminologie suivante. Supposons que X est orientée par une famille cohérente d'orientations $\mathcal{O} = \{\mathcal{O}_{T_x}\}_{x \in X}$ des T_x, $x \in X$; une normale unitaire ξ à X en x s'appelle ***positive*** (par rapport à $\mathcal{O}$) si pour une base $b = \{b_1, \ldots, b_{N-1}\} \in \mathcal{O}_{T_x}$, la base $\{\xi; b\} = \{\xi, b_1, \ldots, b_{N-1}\}$ de $\mathbb{R}^N$ est positive, c'est-à-dire

$$\det[\xi; b] = [\xi, b_1, \ldots, b_{N-1}] > 0 ;$$

si le déterminant est < 0, on dira que ξ est ***négative***. Notons que si $c = \{c_1, \ldots, c_{N-1}\}$ est une autre base dans $\mathcal{O}_{T_x}$ alors $\det[\xi; c] > 0$ si et seulement si $\det[\xi; b] > 0$; en effet, si

$$c_i = \sum_{j=1}^{N-1} t_{ij} b_j , \quad i = 1, \ldots, N-1 ,$$

$$T = [t_{ij}] , \quad M = \begin{bmatrix} 1 & 0 \\ 0 & T \end{bmatrix} ,$$

T une matrice $(N-1, N-1)$, M une matrice (N, N), alors la multilinéarité du déterminant donne que

$$\det[\xi; c] = (\det M) \det[\xi; b] = (\det T) \det[\xi; b]$$

(voir la formule précédant (7.30)) ; l'affirmation est une conséquence du fait que $\det T > 0$.

Si X est orientée par un atlas $\mathcal{A}$ on utilisera la même terminologie ; l'atlas $\mathcal{A}$ orienté correspond à une unique famille cohérente $\mathcal{O} = \{\mathcal{O}_{T_x}\}_{x\in X}$ d'orientations des T_x, $x \in X$; alors une normale unitaire ξ à X en a s'appellera ***positive*** (par rapport à l'orientation $\mathcal{A}$) si elle est positive par rapport à $\mathcal{O}$. Si (ψ, U) est un paramétrage local dans un voisinage de a qui est positivement orienté par rapport à une orientation donnée de l'hypersurface X, la normale unitaire ξ à X en a sera positive si et seulement si

$$\det[\xi, \mathrm{D}_1\psi(u), \ldots, \mathrm{D}_{N-1}\psi(u)] > 0$$

où $u \in U$ est tel que $\psi(u) = a$.

Nous formulons maintenant la proposition concernant l'existence d'un champ normal unitaire continu avec plus de précision.

7.7.5 Proposition

Soit X une hypersurface de classe C^1 dans $\mathbb{R}^N$.

(i) Supposons que X est orienté par une famille cohérente d'orientations $\mathcal{O} = \{\mathcal{O}_{T_x}\}_{x\in X}$; alors il existe un unique champ normal unitaire continu ν tel que chaque normale unitaire $\nu(x)$, $x \in X$, est positive par rapport à $\mathcal{O}$.

(i) Supposons que ν est un champ normal unitaire continu sur X ; alors X possède une unique orientation par rapport à laquelle $\nu(x)$ est positif pour chaque $x \in X$.

DÉMONSTRATION

(i) Soit $b = \{b_1, \ldots, b_{N-1}\} \in \mathcal{O}_{T_x}$; il est clair qu'il existe un unique $\xi \in \mathbb{R}^N$ tel que

$$\xi \cdot b_i = 0\,,\ i = 1, \ldots, N-1\,, \quad \|\xi\| = 1\,, \quad \det[\xi; b] > 0\,.$$

Posons $\nu(x) = \xi$; il reste à voir que $x \mapsto \nu(x)$ est continue. Si $a \in X$ prenons une boule B centrée en a et soit $\theta : B \to \mathbb{R}$ ayant les propriétés signalées au début du paragraphe 7.7.4 ; la fonction $x \mapsto \widetilde{\xi}(x)$ donne une normale unitaire variant continûment sur l'ensemble connexe $B \cap X$; donc $\det[\widetilde{\xi}(x); b]$ est de signe constant dans $B \cap X$; si le signe est positif, $\widetilde{\xi}(x) = \nu(x)$, $x \in B \cap X$; si le signe est négatif, $-\widetilde{\xi}(x) = \nu(x)$, $x \in B \cap X$; dans les deux cas on voit que $x \mapsto \nu(x)$ est continue dans $B \cap X$, en particulier, continue en $x = a$. Le point a dans X étant arbitraire, ceci démontre que ν est l'unique champ normal unitaire continu avec $\nu(x)$ positif par rapport à $\mathcal{O}$ pour tout $x \in X$.

(ii) Soit $a \in X$; prenons un paramétrage local (ψ, U), $\psi = \psi_a$, $U = U_a$, $a \in V_a = V = \psi(U)$; puisque U est connexe, l'expression

$$g_\psi(u) = \det\left[\nu(\psi(u)), \mathrm{D}_1\psi(u), \ldots, \mathrm{D}_{N-1}\psi(u)\right]$$

est de signe constant dans U ; en passant, si nécessaire, à un autre paramétrage local obtenu à partir de $\psi \circ r$ avec

$$r(u_1, u_2, \ldots, u_{N-1}) = (-u_1, u_2, \ldots, u_{N-1})$$

on peut supposer que $g_\psi(u) > 0$, $u \in U$. Démontrons maintenant que l'atlas $\mathcal{A} = \{\psi_a, U_a\}_{a \in X}$ est orienté. Soit (ψ, U), $(\widetilde{\psi}, \widetilde{U})$ deux paramétrages locaux de $\mathcal{A}$ tels que $U \subset \mathbb{R}^{N-1}$, $\widetilde{U} \subset \mathbb{R}^{N-1}$, $\psi : U \to \mathbb{R}^N$, $\widetilde{\psi} : \widetilde{U} \to \mathbb{R}^N$ avec

$$a \in V \cap \widetilde{V} \neq \emptyset, \quad V = \psi(U), \quad \widetilde{V} = \widetilde{\psi}(\widetilde{U}) ;$$

posons

$$\varphi = \widetilde{\psi}^{-1} \circ \psi, \quad W = \psi^{-1}(V \cap \widetilde{V}), \quad \widetilde{W} = \widetilde{\psi}^{-1}(V \cap \widetilde{V}) ;$$

alors $\varphi : W \to \widetilde{W}$, $\widetilde{\psi} \circ \varphi = \psi$ sur W (voir la discussion à propos de la figure 7.1) ; en posant $\varphi(u) = v$, on a pour $u \in W$

$$\widetilde{\psi}(v) = \psi(u), \quad \widetilde{\psi}'(v)\,\varphi'(u) = \psi'(u).$$

Nous savons que

$$g_\psi(u) = \det\left[\nu(\psi(u)), \mathrm{D}_1\psi(u), \ldots, \mathrm{D}_{N-1}\psi(u)\right] > 0, \quad u \in W$$

$$g_{\widetilde{\psi}}(v) = \det\left[\nu(\widetilde{\psi}(v)), \mathrm{D}_1\widetilde{\psi}(v), \ldots, \mathrm{D}_{N-1}\widetilde{\psi}(v)\right] > 0, \quad v \in \widetilde{W}.$$

Supposons que $\psi(\alpha) = a$, $\widetilde{\psi}(\widetilde{\alpha}) = a$, $\alpha \in W$, $\widetilde{\alpha} \in \widetilde{W}$; donc $\varphi(\alpha) = \widetilde{\alpha}$. Puisque $\widetilde{\psi}'(\widetilde{\alpha})\,\varphi'(\alpha) = \psi'(\alpha)$ on a

$$\mathrm{D}_j\psi(\alpha) = \sum_{i=1}^{N-1} t_{ij}\mathrm{D}_i\widetilde{\psi}(\widetilde{\alpha}), \quad j = 1, \ldots, N-1,$$

où $\varphi'(\alpha) = [t_{ij}]$; on en déduit (voir le calcul analogue dans le paragraphe 7.7.4) que

$$g_\psi(\alpha) = \left\{\det\varphi'(\alpha)\right\} \cdot g_{\widetilde{\psi}}(\widetilde{\alpha})$$

d'où $\det\varphi'(\alpha) > 0$. Puisque W est connexe et $\det\varphi'$ est continu non nul dans W, on en déduit que $\det\varphi'(u) > 0$ pour $u \in W$; ceci montre (ψ, U), $(\widetilde{\psi}, \widetilde{U})$ ont la même orientation, établissant ainsi que $\mathcal{A}$ est un atlas orienté. La dernière partie de cette démonstration montre aussi que si $\widetilde{\mathcal{A}} = (\widetilde{\psi}_i, \widetilde{U}_i)_{i \in I}$ est un autre atlas orienté pour X alors $\widetilde{\mathcal{A}}$ et $\mathcal{A}$ donnent la même orientation pour X sous l'hypothèse que, pour tout $x \in X$, $\nu(x)$ soit positif dans cette orientation. ♦

REMARQUE. Comme conséquence de la proposition 7.7.5 on a que les surfaces orientées dans $\mathbb{R}^3$ sont des variétés orientables. Aussi, la frontière régulière $\partial\Omega$ d'une partie ouverte Ω de $\mathbb{R}^N$, munie du champ des normales unitaires dirigées vers l'extérieur, est une hypersurface orientable de $\mathbb{R}^N$. En particulier, si Ω est une partie ouverte de $\mathbb{R}^N$ avec $\operatorname{Fr}\Omega = \partial\Omega$, alors $\operatorname{Fr}\Omega$ est une hypersurface orientable (qui n'est pas nécessairement connexe, même si Ω est connexe)

7.7.6 Orientation d'un bord

Soit X une variété différentielle de dimension $k \geqslant 1$ dans $\mathbb{R}^N$ (de classe C^1). Si A est une partie quelconque de X, nous désignons par $\operatorname{Fr}_X A$, la frontière de A relative à X ; $\operatorname{Fr}_X A$ est formée des points $x \in X$ tels que pour tout voisinage X-ouvert V de x ($x \in V \subset X$), $V \cap A$ et $V \cap A^c$ sont non vides.

Dans ce paragraphe nous considérons les parties Ω de X qui sont X-ouvertes et telles que $\overline{\Omega}$ (adhérence de Ω dans $\mathbb{R}^N$) est contenue dans X ; pour une telle partie Ω, nous écrivons $\operatorname{bord}(\Omega)$ pour $\operatorname{Fr}_X \Omega$. Si $a \in \operatorname{bord}(\Omega)$, nous dirons que a est dans le ***bord régulier*** de Ω (notation : $a \in \partial\Omega$) si la condition suivante est vérifiée : il existe un paramétrage local (ψ, U) avec $\psi : U \to \mathbb{R}^N$, $a \in V = \psi(U) \subset X$, U un domaine de $\mathbb{R}^k$, tel que

$$\psi(U \cap H_k) = V \cap \Omega\,, \quad \psi(U \cap \operatorname{Fr} H_k) = V \cap \operatorname{bord}(\Omega)$$

où $H_k = \{(u_1, \ldots, u_k) \in \mathbb{R}^k : u_1 < 0\}$. Nous dirons que H_k est le ***demi-espace ouvert*** de $\mathbb{R}^k$ formé des $u = (u_1, \ldots, u_k)$ tels que $u_1 < 0$; notons que $\operatorname{Fr} H_k$ est l'hyperplan fermé dans $\mathbb{R}^k$ formé des u tels que $u_1 = 0$.

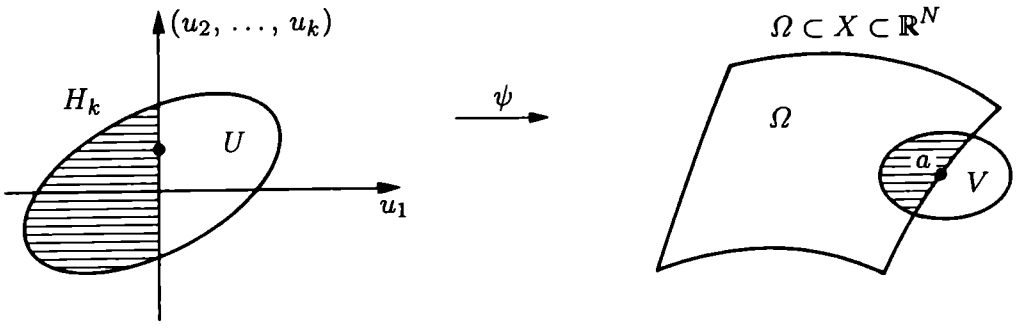


Fig. 7.4

Nous dirons que le paramétrage local (ψ, U) est ***adapté*** à $a \in \partial\Omega$; notons que $V \cap \operatorname{bord}(\Omega)$ est contenu dans $\partial\Omega$ et que (ψ, U) est adapté à tout $x \in V \cap \operatorname{bord}(\Omega) = V \cap \partial\Omega$. Posons

$$S_\Omega = \operatorname{bord}(\Omega) \setminus \partial\Omega\,;$$

S_Ω est la partie non régulière de $\operatorname{bord}(\Omega)$.

REMARQUE. Soit X une partie ouverte non vide de $\mathbb{R}^N$; alors X est une variété différentielle de dimension N dans $\mathbb{R}^N$. Une partie Ω de X est X-ouverte si et seulement si Ω est ouverte dans $\mathbb{R}^N$; aussi, $\text{bord}(\Omega) = \text{Fr}\,\Omega$; la définition de $\partial\Omega$ dans ce cas correspond à ce que nous avons appelé la frontière régulière de Ω dans le paragraphe 7.4.2 (et désigné également par $\partial\Omega$).

Démontrons que *$\partial\Omega$ est une variété différentielle de dimension $(k-1)$ de classe C^1*. Notons que pour $k = 1$ ceci revient à dire que $\partial\Omega$ est un ensemble discret contenu dans X ce qui est évident à partir de la définition ; établissons l'affirmation dans le cas général $k \geqslant 2$. Soit $a \in \partial\Omega$ et (ψ, U) un paramétrage local adapté à $a \in \psi(U) = V$. Posons, pour $(u_1, \ldots, u_{k-1}) \in \mathbb{R}^{k-1}$,

$$\alpha(u_1, \ldots, u_{k-1}) = (0, u_1, \ldots, u_{k-1}) ;$$

alors $\alpha : \mathbb{R}^{k-1} \to \text{Fr}\,H_k$; soit

$$\widehat{\psi}(u_1, \ldots, u_{k-1}) = \psi(0, u_1, \ldots, u_{k-1}) .$$

On a $\widehat{\psi} = \psi \circ \alpha : \widehat{U} \to \widehat{V}$ où

$$\widehat{U} = \alpha^{-1}(U \cap \text{Fr}\,H_k) , \quad \widehat{V} = \widehat{\psi}(\widehat{U})$$

et $a \in \widehat{V} = V \cap \partial\Omega$.

Il est clair que $\widehat{U}$ est un domaine de $\mathbb{R}^{k-1}$ et que $\widehat{\psi}$ est un homéomorphisme de $\widehat{U}$ sur $\widehat{V}$ car ψ et α sont des homéomorphismes ; de plus, $\widehat{\psi}$ est de classe C^1 et

$$\widehat{\psi}'(u_1, \ldots, u_{k-1}) = \begin{bmatrix} \mathrm{D}_2\psi_1 & & \mathrm{D}_k\psi_1 \\ \vdots & & \vdots \\ \mathrm{D}_2\psi_k & & \mathrm{D}_k\psi_k \end{bmatrix} (0, u_1, \ldots, u_{k-1})$$

d'où $\text{rang}\,\widehat{\psi}' = (k-1)$ partout dans $\widehat{U}$ car $\text{rang}\,\psi' = k$ partout dans U. Ceci démontre que $(\widehat{\psi}, \widehat{U})$ est un paramétrage local pour le voisinage $\widehat{V}$ de $a \in \partial\Omega$, établissant que $\partial\Omega$ est une variété différentielle de dimension $(k-1)$ de classe C^1. ♦

De ce qui précède l'on voit que si

$$\psi(0, u_1, \ldots, u_{k-1}) = a$$

où $(0, u_1, \ldots, u_{k-1}) \in U \cap \text{Fr}\,H_k$, alors $T_a(\partial\Omega)$ est le sous-espace de $\mathbb{R}^N$ engendré par

$$\{\mathrm{D}_2\psi(0; u'), \ldots, \mathrm{D}_k\psi(0; u')\}$$

où $u' = (u_1, \ldots, u_{k-1})$, $(0; u') = (0, u_1, \ldots, u_{k-1})$. Rappelons que $T_a(X)$ est le sous-espace de $\mathbb{R}^N$ engendré par

$$\{\mathrm{D}_1\psi(0; u'), \mathrm{D}_2\psi(0; u'), \ldots, \mathrm{D}_k\psi(0; u')\} .$$

Supposons maintenant que X est orientable et que son orientation est donnée par une famille cohérente d'orientations $\mathcal{O} = \{\mathcal{O}_{T_x(X)}\}_{x\in X}$. Posons

$$\widehat{\mathcal{O}} = \{\widehat{\mathcal{O}}_{T_a(\partial\Omega)}\}_{a\in\partial\Omega}$$

où $\widehat{\mathcal{O}}_{T_a(\partial\Omega)}$ est défini comme suit : soit (ψ, U) un paramétrage local d'un voisinage V de a adapté à $a \in \partial\Omega$ tel que $\psi(0; u') = a$ et

$$\{\mathrm{D}_1\psi(0;u'), \mathrm{D}_2\psi(0;u'), \ldots, \mathrm{D}_k\psi(0;u')\} \in \mathcal{O}_{T_a(X)} ;$$

nous orientons $T_a(\partial\Omega)$ en demandant que

$$\{\mathrm{D}_2\psi(0;u'), \ldots, \mathrm{D}_k\psi(0;u')\} \in \widehat{\mathcal{O}}_{T_a(\partial\Omega)} .$$

Il est clair de cette définition que $\widehat{\mathcal{O}}$ est une famille cohérente d'orientations des espaces tangents de $\partial\Omega$. L'orientation de $\partial\Omega$ ainsi obtenue de l'orientation de X s'appellera l'***orientation induite*** de $\partial\Omega$.

Notons un cas particulièrement simple mais important. Si X est une variété compacte orientée dans $\mathbb{R}^N$, de dimension $k < N$, alors $\Omega = X$ est une partie X-ouverte avec $\overline{\Omega} = X$ et on vérifie que

$$\mathrm{Fr}_X\, \Omega = \mathrm{bord}(\Omega) = \partial\Omega = \emptyset ;$$

dans ce cas, la question d'orientation de $\partial\Omega$ ne se pose pas.

REMARQUE. L'orientation induite du bord régulier $\partial\Omega$ est uniquement déterminée par l'orientation de la variété ambiante X, Ω étant une partie X-ouverte dans $X \subset \mathbb{R}^N$ avec $\overline{\Omega} \subset X$ ($\overline{\Omega}$ étant l'adhérence de Ω dans $\mathbb{R}^N$). Il y a d'autres manières d'obtenir une orientation de $\partial\Omega$; celle que nous appelons « l'orientation induite » de $\partial\Omega$ est précisément l'orientation nécessaire pour la validité du théorème de Stokes généralisé.

7.7.7 Compléments

Une façon élégante de traiter la théorie des formes différentielles et l'orientation dans une variété X de dimension k de classe C^r est d'introduire les ***fibrés vectoriels*** sur X, en particulier, le ***fibré tangent*** $T(X)$ et le ***fibré des p-formes alternées*** $\mathcal{A}^p(X)$. Comme ensembles, on a

$$T(X) = \bigcup_{x\in X} [\{x\} \times T_x(X)] \subset \mathbb{R}^N \times \mathbb{R}^N$$

$$\mathcal{A}^p(X) = \bigcup_{x\in X} [\{x\} \times \mathcal{A}_p(T_x(X))] \subset \mathbb{R}^N \times \mathbb{R}^{\binom{N}{p}} .$$

Ces ensembles sont considérés comme des variétés différentielles de dimension $2k$ et $k + \binom{k}{p}$ respectivement. Un ***champ vectoriel*** sur X est une ***section*** s

de $T(X)$: $s(x) \in T_x(X)$ pour $x \in X$; une p-forme différentielle sur X est une section ω de $\mathcal{A}^p(X)$:

$$\omega(x) \in \mathcal{A}_p\big(T_x(X)\big)\,, \quad x \in X\,.$$

Une k-forme continue sur X de dimension k est déterminée par une fonction continue $f : X \to \mathbb{R}$ dans le sens que

$$\omega(x) = f(x)u_1(x) \wedge \cdots \wedge u_k(x)\,, \quad x \in X\,,$$

où $\{u_1(x), \ldots, u_k(x)\}$ est une base pour $\mathcal{A}_1\big(T_x(X)\big) = T'_x(X)$ qui varie continûment avec x. On dit que X est orientable s'il existe une k-forme *continue* ω partout non nulle, c'est-à-dire $f(x) \neq 0$, $x \in X$. On voit que cette définition d'orientabilité est liée à la notion d'orientation d'un espace vectoriel de dimension k (en l'occurrence, $T_x(X)$) par le choix d'une k-forme alternée non nulle ; ensuite il s'agit de varier ces orientations de $T_x(X)$ continûment avec x ce qui correspond à l'idée d'avoir une famille cohérente d'orientations introduite dans le paragraphe 7.7.3.

Le développement esquissé ici est donné en détail dans Dieudonné [13, tome III, pp. 74-153]. Le travail technique nécessaire pour faire ce développement est amplement compensé par les multitudes d'applications qui deviennent possibles par la suite.

7.8 INTÉGRATION DES FORMES DIFFÉRENTIELLES

7.8.1 Introduction

Soit G une partie ouverte non vide de $\mathbb{R}^N$. Dans ce paragraphe nous définissons

$$\int_A \omega$$

lorsque ω est une k-forme dans G et A est une partie d'une variété différentielle *orientée* X de dimension k (de classe C^1) contenue dans G. Pour abréger, on dira simplement que X est une variété orientée de dimension k dans G ; en général, ω sera une k-forme continue ; A sera *toujours* une partie borélienne de X et cette condition ne sera pas mentionnée explicitement.

7.8.2 Intégration d'une N-forme

Une N-forme ω dans $G \subset \mathbb{R}^N$ (G étant ouvert non vide) possède une écriture canonique

$$\omega = f\,\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_N$$

où $f : G \to \mathbb{R}$; on supposera *toujours* que f est une fonction borrélienne ; f est continue si ω est continue (d'après la définition de la continuité ou de l'appartenance à la classe C^r d'une forme ω donnée dans le paragraphe 7.6.2).

Si $A \subset G$ on dira que ω est ***intégrable*** dans A si

$$\int_A |f(x)|\,\mathrm{d}x < \infty\,;$$

dans ce cas on pose

$$\int_A \omega = \int_A f(x)\,\mathrm{d}x\,.$$

Ainsi, l'intégrale d'une N-forme dans G se ramène à une intégrale ordinaire dans $\mathbb{R}^N$.

Soit $\varphi : G \to H$ une application *injective* de classe C^1 de G dans une partie ouverte H de $\mathbb{R}^N$; si ω est une N-forme dans H, $\varphi^*\omega$ sera une N-forme dans G. Notons les fonctions coordonnées dans H comme $x_1, \ldots, x_N$ et celles dans G comme $u_1, \ldots, u_N$; si

$$\omega = f\,\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_N\,, \quad f : H \to \mathbb{R},$$

on aura

$$\varphi^*\omega = (f \circ \varphi)\,\mathrm{d}\varphi_1 \wedge \cdots \wedge \mathrm{d}\varphi_N = (f \circ \varphi)\det\varphi'\,\mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_N$$

où $\varphi = (\varphi_1, \ldots, \varphi_N)$ et

$$\det\varphi' = \frac{\partial(\varphi_1, \ldots, \varphi_N)}{\partial(u_1, \ldots, u_N)}\,.$$

Si $\det\varphi' > 0$ *partout dans* G (on dit que φ ***préserve les orientations***) *et si* ω *est intégrable dans* $A \subset H$, *alors* $\varphi^*\omega$ *est intégrable dans* $\varphi^{-1}(A)$ *et*

$$\int_A \omega = \int_{\varphi^{-1}(A)} \varphi^*\omega\,.$$

En effet, d'après la définition donnée ci-dessus,

$$\begin{aligned}\int_{\varphi^{-1}(A)} \varphi^*\omega &= \int_{\varphi^{-1}(A)} f\big(\varphi(u)\big)\det\varphi'(u)\,\mathrm{d}u \\ &= \int_A f(x)\,\mathrm{d}x = \int_A \omega\end{aligned}$$

en utilisant la formule de changement de variables dans les intégrales donnée dans la section 5.7 car ici $|\det\varphi'| = \det\varphi'$ à cause de l'hypothèse $\det\varphi' > 0$.

D'autre part, si $\det\varphi' < 0$ *partout dans* G (on dit que φ ***renverse les orientations***) *et si* ω *est intégrable dans* $A \subset H$ *alors* $\varphi^*\omega$ *est intégrable dans* $\varphi^{-1}(A)$ *et*

$$\int_A \omega = -\int_{\varphi^{-1}(A)} \varphi^*\omega\,.$$

Ceci est une conséquence immédiate du calcul précédent car dans ce cas $|\det\varphi'| = -\det\varphi'$ à cause de l'hypothèse $\det\varphi' < 0$.

Notons que si G est connexe et $\varphi : G \to H$ est un C^1-difféomorphisme alors $\det \varphi' \neq 0$ partout dans G ; alors on aura ou bien $\det \varphi' > 0$ partout dans G ou bien $\det \varphi' < 0$ partout dans G.

7.8.3 Intégration d'une k-forme

Soit ω une k-forme dans G (G étant ouvert non vide dans $\mathbb{R}^N$) avec $1 \leqslant k < N$. Soit X une variété *orientée* de dimension k dans G ; pour la discussion qui suit on travaille avec cette orientation fixée dans X.

Soit $A \subset X$; supposons d'abord que A est *contenu dans un morceau* $V = \psi(U)$ où $\psi : U \to \mathbb{R}^N$ est un paramétrage *positivement orienté* ; nous dirons que ω est ***intégrable*** dans A si $\psi^*\omega$ est intégrable dans $\psi^{-1}(A) \subset U$. Noter que $\psi^*\omega$ est une k-forme dans U, une partie ouverte (en fait, un domaine) de $\mathbb{R}^k$; donc l'intégrale de $\psi^*\omega$ sur les parties de U entre dans le cadre du paragraphe 7.8.2. On pose (pour $\psi^*\omega$ intégrable dans A) :

$$\int_A \omega = \int_{\psi^{-1}(A)} \psi^*\omega \,.$$

Vérifions que *ni l'intégrabilité de ω dans A ni la valeur de $\int_A \omega$ ne dépend du paramétrage (ψ, U) utilisé.* En effet, soit $(\widetilde{\psi}, \widetilde{U})$ un autre paramétrage *positivement orienté* tel que

$$\widetilde{V} = \widetilde{\psi}(\widetilde{U}) \supset A \,;$$

posons (cf. fig. 7.1) :

$$W = \psi^{-1}(V \cap \widetilde{V}), \quad \widetilde{W} = \widetilde{\psi}^{-1}(V \cap \widetilde{V}), \quad \varphi = \widetilde{\psi}^{-1} \circ \psi$$

avec $\varphi : W \to \widetilde{W}$; rappelons que $\widetilde{U}$ est un domaine de $\mathbb{R}^k$ (ainsi que U), $\widetilde{W}$ est un sous-domaine de $\widetilde{U}$, W est un sous-domaine de U et $\varphi : W \to \widetilde{W}$ est un C^1-difféomorphisme tel que $\det \varphi' > 0$ dans W. Aussi

$$\psi^{-1}(A) \subset W, \quad \widetilde{\psi}^{-1}(A) \subset \widetilde{W}, \quad \varphi^{-1}\big(\widetilde{\psi}^{-1}(A)\big) = \psi^{-1}(A)$$

car $\widetilde{\psi} \circ \varphi = \psi$; à l'aide du paragrphe 7.8.2 on obtient que

$$\int_{\widetilde{\psi}^{-1}(A)} \widetilde{\psi}^*\omega = \int_{\varphi^{-1}(\widetilde{\psi}^{-1}(A))} \varphi^*(\widetilde{\psi}^*\omega) = \int_{\psi^{-1}(A)} \psi^*\omega$$

car d'après le paragraphe 7.6.5 on a

$$\varphi^*(\widetilde{\psi}^*\omega) = \big(\widetilde{\psi} \circ \varphi\big)^*\omega = \psi^*\omega \,;$$

ceci établit l'indépendance de $\int_A \omega$ du paramétrage choisi pour sa définition.

Supposons que ω est une k-forme dans G telle que $\operatorname{supp}\omega$ est contenu dans une partie ouverte $B \subset G$ telle que $B \cap X = V$ soit un morceau ; par $\operatorname{supp}\omega$ nous entendons l'adhérence des points $x \in G$ tels que $\omega(x) \neq 0$. Dans ce

cas, on dira que ω est ***intégrable*** dans $A \subset X$ si ω est intégrable dans $A \cap V$ et on pose

$$\int_A \omega = \int_{A \cap V} \omega$$

où le second membre est défini comme ci-dessus par un paramétrage positivement orienté $\psi : U \to \mathbb{R}^N$, $\psi(U) = V$.

Il est clair que

$$\omega \longmapsto \int_A \omega$$

est linéaire pour les formes linéaires intégrables dans A envisagées ci-dessus. Pour définir l'intégrabilité de ω dans A et la valeur de $\int_A \omega$ pour une partie (borélienne) quelconque A de X nous utilisons une partition de l'unité exactement comme dans le paragraphe 7.3.7. Esquissons le raisonnement : pour $x \in X$ soit $V_x = \psi_x(U_x)$ un morceau contenant x où $V_x = B_x \cap X$, B_x étant un voisinage ouvert de x dans $\mathbb{R}^N$ contenu dans G et $\psi_x : U_x \to \mathbb{R}^N$, $U_x \subset \mathbb{R}^k$, un paramétrage *positivement orienté* ; posons $\Omega = \bigcup_{x \in X} B_x$, $\Omega \subset G$. Soit $\{\theta_i\}$ une partition de l'unité subordonnée à $\{B_x\}_{x \in X}$ avec $\operatorname{supp} \theta_i \subset B_{x_i} = B_i$; posons $\psi_{x_i} = \psi_i$, $U_{x_i} = U_i$, $V_{x_i} = V_i$. Notons que $\operatorname{supp} \theta_i \omega$ est compact et

$$\operatorname{supp}(\theta_i \omega) \subset B_i \ ;$$

en particulier, $\theta_i \omega = 0$ en dehors du morceau V_i ; de plus,

$$\omega = \sum_{i \geqslant 1} \theta_i \omega$$

dans Ω car $\sum_{i \geqslant 1} \theta_i = 1$ sur Ω. Nous dirons que ω est ***intégrable*** dans $A \subset X$ si $\theta_i \omega$ est intégrable dans A pour chaque i et

$$\sum_{i \geqslant 1} \left| \int_{A \cap V_i} \theta_i \omega \right| = \sum_{i \geqslant 1} \left| \int_A \theta_i \omega \right| < \infty \, .$$

Si ω est intégrable dans A, nous définissons

$$\int_A \omega = \sum_{i \geqslant 1} \int_{A \cap V_i} \theta_i \omega = \sum_{i \geqslant 1} \int_A \theta_i \omega \ ; \tag{7.34}$$

en écrivant $A_i = A \cap V_i$, on aura

$$\int_{A_i} \theta_i \omega = \int_{\psi_i^{-1}(A_i)} \psi_i^*(\theta_i \omega) = \int_{\psi_i^{-1}(A_i)} (\theta_i \circ \psi_i) \psi_i^* \omega \, .$$

En utilisant le raisonnement du paragraphe 7.3.7, on montre sans peine que ni l'intégrabilité de ω dans A ni la valeur de $\int_A \omega$ ne dépendent de la partition de l'unité utilisée.

Il est clair que l'ensemble des k-formes intégrables dans $A \subset X$ est un espace vectoriel et que l'application

$$\omega \longmapsto \int_A \omega$$

est une forme linéaire sur cet espace vectoriel. De plus, si ω est une k-forme intégrable dans les parties A et B de X alors ω est intégrable dans $A \cup B$; si A et B sont disjoints, on a aussi

$$\int_{A\cup B} \omega = \int_A \omega + \int_B \omega .$$

Ces considérations se généralisent facilement à un nombre fini quelconque de parties de X.

7.8.4 Une inégalité pour les intégrales d'une k-forme

Soit ω une k-forme dans une partie ouverte G de $\mathbb{R}^N$; soit X une variété différentielle *orientée* de dimension $k \geqslant 1$ dans G (de classe C^1).

Supposons que l'écriture canonique de ω est

$$\omega = \sum_{I\in S(k;N)} f_I \, \mathrm{d}x_I .$$

Soit $\psi : U \to \mathbb{R}^N$, $U \subset \mathbb{R}^k$, $\psi(U) \subset X$ où (ψ, U) est un paramétrage positivement orienté (par rapport à l'orientation fixée de X). On a (avec $I = \{i_1 < \ldots < i_k\}$) :

$$\psi^*\omega = \sum_{I\in S(k;N)} (f_I \circ \psi)\, \mathrm{d}\psi_{i_1} \wedge \cdots \wedge \mathrm{d}\psi_{i_k} = \sum_I (f_I \circ \psi) \cdot J_I^\psi \cdot (\mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_k)$$

où (cf. (7.33), § 7.6.5) :

$$J_I^\psi = \frac{\partial(\psi_{i_1}, \ldots, \psi_{i_k})}{\partial(u_1, \ldots, u_k)} .$$

Notons que

$$\rho_\psi = \left\{ \sum_I |J_I^\psi|^2 \right\}^{1/2}$$

donne l'élément de volume dans X :

$$\mathrm{d}\sigma^X = \rho_\psi \, \mathrm{d}u$$

((7.7), § 7.3.6). Posons

$$g = \left\{ \sum_I |f_I|^2 \right\}^{1/2} ; \tag{7.35}$$

de l'inégalité de Cauchy-Schwarz on a

$$\left|\sum_I (f_I \circ \psi) J_I^\psi\right| \leqslant \left\{\sum_I |f_I \circ \psi|^2\right\}^{1/2} \left\{\sum_I |J_I^\psi|^2\right\}^{1/2} = (g \circ \psi)\, \rho_\psi .$$

Si $A \subset \psi(U) \subset X$ et

$$\int_A g \, d\sigma^X = \int_{\psi^{-1}(A)} (g \circ \psi)\, \rho_\psi \, du < \infty$$

alors ω sera intégrable dans A et

$$\left|\int_A \omega\right| \leqslant \int_A g \, d\sigma^X \tag{7.36}$$

car

$$\left|\int_A \omega\right| = \left|\int_{\psi^{-1}(A)} \sum_I f_I \circ \psi \cdot J_I^\psi \, du\right| \leqslant \int_{\psi^{-1}(A)} (g \circ \psi)\, \rho_\psi \, du = \int_A g \, d\sigma^X .$$

Nous démontrons maintenant (7.36) en toute généralité. *Si $A \subset X$ et $\int_A g \, d\sigma^\psi < \infty$ où g est défini par (7.35) alors $\omega = \sum_I f_I \, dx_I$ est intégrable dans A et (7.36) est juste.*

En effet, en utilisant (7.34), on a

$$\left|\int_A \omega\right| \leqslant \sum_{i \geqslant 1} \left|\int_{A \cap V_i} \theta_i \omega\right| \leqslant \sum_{i \geqslant 1} \int_A g\, \theta_i \, d\sigma^X = \int_A g \, d\sigma^X$$

en se rappelant que $0 \leqslant \theta_i \leqslant 1$ et $\sum_i \theta_i = 1$ sur X. ♦

Notons un corollaire immédiat de (7.36). Si

$$\omega = \sum_I f_I \, dx_I ,$$

$$M' = \sup_{I,\, x \in A} |f_I(x)| , \quad M = M' \cdot \binom{N}{k} < \infty$$

et $\sigma^X(A) < \infty$ alors ω est intégrable dans $A \subset X$ et

$$\left|\int_A \omega\right| \leqslant M\, \sigma^X(A) . \tag{7.37}$$

L'inégalité (7.37) est applicable en particulier si A est une partie compacte de X et ω est continue.

7.8.5 Intégration sur une hypersurface orientée

Soit X une hypersurface orientée (de classe C^1) dans $\mathbb{R}^N$; nous supposons que l'orientation de X est fixée par le choix d'un champ normal unitaire continu ν sur X (prop. 7.7.5).

Soit $F = (F_1, \ldots, F_N)$ un champ vectoriel dans une partie ouverte G de $\mathbb{R}^N$ avec $G \supset X$; à F nous associons la $(N-1)$-forme différentielle ω dans G donnée par

$$\omega = \sum_{i=1}^{N} (-1)^{i-1} F_i \, \mathrm{d}x_1 \wedge \cdots \wedge \widehat{\mathrm{d}x_i} \wedge \cdots \wedge \mathrm{d}x_N \,,$$

le signe $\widehat{\ }$ indiquant que le terme concerné est absent. Nous écrivons σ^X pour la mesure naturelle de X.

Ces notations étant fixées, nous pouvons énoncer ce qui suit : *la forme ω est intégrable dans une partie A de X si et seulement si $F \cdot \nu$ est σ^X-intégrable dans A et alors*

$$\int_A \omega = \int_A F \cdot \nu \, \mathrm{d}\sigma^X \tag{7.38}$$

Notons pour l'usage dans la suite que $\mathrm{d}\omega = (\operatorname{div} F)\, \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_N$ où $\operatorname{div} F = \mathrm{D}_1 F_1 + \cdots + \mathrm{D}_N F_N$ (Ex. 7, § 7.1.7).

DÉMONSTRATION. Supposons d'abord que $A \subset \psi(U)$ où (ψ, U) est un paramétrage positivement orienté dans X, U étant un domaine dans $\mathbb{R}^{N-1}$ avec $\psi = (\psi_1, \ldots, \psi_{N-1})$. Alors

$$\begin{aligned} \psi^* \omega &= \sum_{i=1}^{N} (-1)^{i-1} F_i \circ \psi \, \mathrm{d}\psi_1 \wedge \cdots \wedge \widehat{\mathrm{d}\psi_i} \wedge \cdots \wedge \mathrm{d}\psi_N \\ &= \det\bigl[F \circ \psi, \mathrm{D}_1\psi, \ldots, \mathrm{D}_{N-1}\psi\bigr] \, \mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_{N-1} \end{aligned} \tag{7.39}$$

en notant les fonctions coordonnées dans U par $u_1, \ldots, u_{N-1}$ (voir (7.33) et développer le déterminant ci-dessus suivant sa première colonne).

Puisque (ψ, U) est un paramétrage positivement orienté on a que

$$\det\bigl[\nu\bigl(\psi(u)\bigr), \mathrm{D}_1\psi(u), \ldots, \mathrm{D}_{N-1}\psi(u)\bigr] > 0 \,, \quad u \in U \,,$$

car l'orientation de $T_{\psi(u)}(X)$ est fixée par la base $\{\mathrm{D}_1\psi(u), \ldots, \mathrm{D}_{N-1}\psi(u)\}$ (§ 7.7.4).

Si $\xi_1, \ldots, \xi_{N-1}$ sont $(N-1)$ vecteurs linéairement indépendants dans $\mathbb{R}^N$, posons

$$u(a) = \det\bigl[a, \xi_1, , \ldots, \xi_{N-1}\bigr] \,, \quad a \in \mathbb{R}^N \,.$$

Puisque $a \mapsto u(a)$ est linéaire et u n'est pas identiquement nulle, il existe un unique $\xi = \phi(\xi_1, \ldots, \xi_{N-1})$ dans $\mathbb{R}^N$ tel que

$$a \cdot \xi = u(a) = \det\bigl[a, \xi_1, , \ldots, \xi_{N-1}\bigr] \,, \quad a \in \mathbb{R}^N \,;$$

il est clair que ξ satisfait aux conditions suivantes :

$$\xi \cdot \xi_i = 0\,,\ 1 \leqslant i \leqslant N-1\,, \quad \det[\xi, \xi_1, ,\ldots, \xi_{N-1}] = \|\xi\|^2 > 0\,.$$

Soit M la matrice $(N-1\,,\ N-1)$ définie par $[\xi_1, ,\ldots, \xi_{N-1}]$; alors

$$\begin{aligned}\det{}^2[\xi, \xi_1, ,\ldots, \xi_{N-1}] &= \det\{{}^t[\xi, \xi_1, ,\ldots, \xi_{N-1}]\,[\xi, \xi_1, ,\ldots, \xi_{N-1}]\} \\ &= \det\begin{bmatrix} \|\xi\|^2 & 0 \\ 0 & {}^tMM \end{bmatrix} = \|\xi\|^2 \det({}^tMM)\end{aligned}$$

d'où la relation

$$\|\xi\|^4 = \|\xi\|^2 \det({}^tMM)$$

donne que

$$\|\xi\| = \{\det({}^tMM)\}^{1/2}\,.$$

Utilisons ce qui précède avec $\xi_i = \mathrm{D}_i\psi(u)$, $1 \leqslant i \leqslant N-1$; du fait que

$$\nu(\psi(u)) \cdot \mathrm{D}_i\psi(u) = 0\,, \quad 1 \leqslant i \leqslant N-1\,,$$

et

$$\det[\nu \circ \psi, \mathrm{D}_1\psi, \ldots, \mathrm{D}_{N-1}\psi] > 0\,,$$

on déduit que, si $\xi(u) = \phi(\mathrm{D}_1\psi(u), \ldots, \mathrm{D}_{N-1}\psi(u))$, alors

$$\nu(\psi(u)) = \frac{\xi(u)}{\|\xi(u)\|}$$

et

$$\begin{aligned}F(\psi(u)) \cdot \nu(\psi(u)) &= F(\psi(u)) \cdot \frac{\xi(u)}{\|\xi(u)\|} \\ &= \det[F(\psi(u)), \mathrm{D}_1\psi(u), \ldots, \mathrm{D}_{N-1}\psi(u)] \cdot \frac{1}{\|\xi(u)\|}\,;\end{aligned}$$

puique ici M est égal à $\psi'(u)$, on a

$$\|\xi(u)\| = \{\det({}^t\psi'(u)\psi'(u))\}^{1/2} = \rho_\psi(u)$$

où $\rho_\psi(u)$ donne l'élément du volume dans X, c'est-à-dire

$$\mathrm{d}\sigma^X = \rho_\psi\,\mathrm{d}u\,.$$

Donc,

$$F \cdot \nu\,\mathrm{d}\sigma^X = \det[F \circ \psi, \mathrm{D}_1\psi, \ldots, \mathrm{D}_{N-1}\psi]\,\mathrm{d}u\,. \tag{7.40}$$

De (7.39) et (7.40) on déduit (7.38) ainsi que les affirmations concernant l'intégrabilité car les deux membres de (7.38) correspondent à l'expression

$$\int_{\psi^{-1}(A)} \det[F \circ \psi, \mathrm{D}_1\psi, \ldots, \mathrm{D}_{N-1}\psi]\,\mathrm{d}u\,.$$

Le cas général d'un ensemble $A \subset X$ est établi à partir du cas particulier précédent en utilisant une partition de l'unité. ♦

7.8.6 Remarques

Notons que contrairement à l'intégration des fonctions scalaires sur une variété X (§ 7.3.7) on ne peut pas définir l'intégrale d'une k-forme ω sur une variété X de dimension k sans que X soit orientée. En effet, la valeur $\int_A \omega$ lorsque A est contenu dans un morceau paramétré par (ψ, U), change de signe si l'on passe à un autre paramétrage $(\widetilde{\psi}, \widetilde{U})$ qui a une orientation opposée ; ainsi, pour pouvoir définir $\int_A \omega$ pour une partie générale A de X, il est nécessaire d'avoir un atlas orienté de X (c'est-à-dire, une famille de paramétrages couvrant tout X et ayant les mêmes orientations).

Supposons que ω est une k-forme intégrable dans $A \subset X$, X étant une variété orientée de dimension k ; soit $\mathcal{A}$ l'atlas qui donne l'orientation de X. On sait (§ 7.7.1) que X possède une orientation opposée donnée par un atlas $\widetilde{\mathcal{A}}$. Il est facile à démontrer que ω est aussi intégrable dans $A \subset X$, X étant orientée par $\widetilde{\mathcal{A}}$; en notant par $\widetilde{\int_A} \omega$ l'intégrale de ω sur A où A est considéré comme une partie de X orientée par $\widetilde{\mathcal{A}}$, on a

$$\widetilde{\int_A} \omega = -\int_A \omega . \tag{7.41}$$

Il est utile de voir ce que (7.41) exprime dans les deux cas où $k = 1$ et $k = (N - 1)$. Une 1-forme est intégrée sur une variété X de dimension 1 ; si X est connexe, alors X est une courbe simple régulière qui a exactement deux orientations données par les deux sens de parcours de la courbe ; dans ce cas, (7.41) correspond à une formule bien connue de la théorie des intégrales curvilignes. Si $k = (N-1)$, X est une hypersurface orientée dans $\mathbb{R}^N$ et l'orientation opposée correspond au choix d'un champ normal unitaire continu du signe opposé ; dans ce cas (7.41) est évidente à partir de (7.38).

7.8.7 Une formule utile

Nous reprenons d'abord quelques notations du paragraphe 7.7.6. Soit H_k le demi-espace ouvert de $\mathbb{R}^k$ formé des $u = (u_1, \ldots, u_k)$ tels que $u_1 < 0$; nous écrivons ∂H_k pour $\operatorname{Fr} H_k$; cette écriture concorde bien avec le fait que $\operatorname{Fr} H_k$ est le bord régulier de la partie ouverte H_k de la variété $\mathbb{R}^k$. Nous considérons $\mathbb{R}^k$ comme une variété orientée, dotée de son orientation usuelle ; alors l'orientation induite dans ∂H_k, une hypersurface dans $\mathbb{R}^k$, est celle fixée par le choix du champ normal unitaire constant ν défini comme suit :

$$\nu(x) = e_1 = (1, 0, \ldots, 0), \quad x \in \partial H_k .$$

En effet, à chaque point $a \in \partial H_k$, $\mathcal{O}_{T_a(\partial H_k)}$ est fixée en prenant $\{e_2, \ldots, e_k\}$ comme une base ordonnée positive ; l'unique vecteur unitaire ν orthogonal aux $e_2, \ldots, e_k$ et tel que $\{\nu, e_2, \ldots, e_k\}$ est une base ordonnée positive pour $T_a(\mathbb{R}^k)$ est $\nu = e_1$.

Ces rappels étant faits, nous énonçons la formule : *si ω est une $(k-1)$-forme de classe C^1 à support compact dans $\mathbb{R}^k$, $k \geqslant 2$, alors*

$$\int_{H_k} \mathrm{d}\omega = \int_{\partial H_k} \omega \,. \tag{7.42}$$

REMARQUE. Notons que (7.42) est vrai pour $k = 1$ aussi pourvu que l'on interprète l'orientation de $\partial H_1 = \{0\}$ et l'intégration sur $\{0\}$ convenablement. En effet,

$$H_1 = \,]-\infty, 0[\, \subset \mathbb{R}^1 \,;$$

une 0-forme ω de classe C^1 à support compact dans $\mathbb{R}^1 = \mathbb{R}$ est une fonction $f : \mathbb{R} \to \mathbb{R}$ de classe C^1 avec $f(x) = 0$ si $|x| > T$ pour un $T > 0$; alors $\mathrm{d}\omega = f'(x)\,\mathrm{d}x$ et

$$\int_{H_1} \mathrm{d}\omega = \int_{-\infty}^{0} f'(x)\,\mathrm{d}x = f(0)\,.$$

Il suffit maintenant d'interpréter l'intégrale de $f = \omega$ sur $\{0\} = \partial H_1$ sous l'orientation induite comme étant $f(0)$ pour avoir la validité de (7.42) pour $k = 1$. Cette discussion montre aussi la relation entre (7.42) et le théorème fondamental du calcul différentiel et intégral.

DÉMONSTRATION DE (7.42). Supposons que l'écriture canonique de ω est

$$\omega = \sum_{i=1}^{k} f_i \,\mathrm{d}x_1 \wedge \cdots \wedge \widehat{\mathrm{d}x_i} \wedge \cdots \wedge \mathrm{d}x_k$$

où $\widehat{\ }$ indique l'absence du terme concerné ; les fonctions $f_i : \mathbb{R}^k \to \mathbb{R}$ sont de classe C^1 à supports compacts. Notons que

$$\begin{aligned}
\mathrm{d}\omega &= \sum_{i=1}^{k} \mathrm{d}f_i \wedge \mathrm{d}x_1 \wedge \cdots \wedge \widehat{\mathrm{d}x_i} \wedge \cdots \wedge \mathrm{d}x_k \\
&= \sum_{i=1}^{k} \mathrm{D}_i f_i \,\mathrm{d}x_i \wedge \mathrm{d}x_1 \wedge \cdots \wedge \widehat{\mathrm{d}x_i} \wedge \cdots \wedge \mathrm{d}x_k \\
&= \sum_{i=1}^{k} (-1)^{i-1} \mathrm{D}_i f_i \,\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_k \,.
\end{aligned}$$

Donc

$$\int_{H_k} \mathrm{d}\omega = \sum_{i=1}^{k} (-1)^{i-1} \int_{H_k} \mathrm{D}_i f_i \,\mathrm{d}x = \int_{\mathbb{R}^{k-1}} f_1(0, x_2, \ldots, x_k)\,\mathrm{d}x_2 \cdots \mathrm{d}x_k \,, \tag{7.43}$$

car pour $i \geqslant 2$

$$\int_{H_k} \mathrm{D}_i f_i \, \mathrm{d}x = \int_{]-\infty,0[\times\mathbb{R}^{k-2}} \left\{ \int_{\mathbb{R}} \mathrm{D}_i f_i \, \mathrm{d}x_i \right\} \mathrm{d}x_1 \cdots \widehat{\mathrm{d}x_i} \cdots \mathrm{d}x_k = 0$$

à cause du fait que

$$\int_{-\infty}^{\infty} \mathrm{D}_i f_i \, \mathrm{d}x_i = 0 \,,$$

f_i étant à support compact ; pour $i = 1$, on a

$$\begin{aligned} \int_{H_k} \mathrm{D}_1 f_1 \, \mathrm{d}x &= \int_{\mathbb{R}^{k-1}} \left\{ \int_{-\infty}^{0} \mathrm{D}_1 f_1 \, \mathrm{d}x_1 \right\} \mathrm{d}x_2 \cdots \mathrm{d}x_k \\ &= \int_{\mathbb{R}^{k-1}} f_1(0, x_2, \ldots, x_k) \, \mathrm{d}x_2 \cdots \mathrm{d}x_k \,. \end{aligned}$$

Pour calculer le second membre de (7.42) nous utilisons le paramétrage suivant pour ∂H_k qui correspond à l'orientation induite de ∂H_k :

$$\alpha(u_1, \ldots, u_{k-1}) = (0, u_1, \ldots, u_{k-1}) \,, \quad (u_1, \ldots, u_{k-1}) \in \mathbb{R}^{k-1} \,.$$

Ici $\alpha = (\alpha_1, \ldots, \alpha_k)$ avec $\alpha_1(u_1, \ldots, u_{k-1}) = 0$, $\alpha_2(u_1, \ldots, u_{k-1}) = u_1$, ..., $\alpha_k(u_1, \ldots, u_{k-1}) = u_{k-1}$; donc

$$\begin{aligned} \alpha^* \omega &= \sum_{i=1}^{k} (f_i \circ \alpha) \, \mathrm{d}\alpha_1 \wedge \cdots \wedge \widehat{\mathrm{d}\alpha_i} \wedge \cdots \wedge \mathrm{d}\alpha_k \\ &= (f_1 \circ \alpha) \, \mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_{k-1} \end{aligned}$$

et

$$\begin{aligned} \int_{\partial H_k} \omega &= \int_{\mathbb{R}^{k-1}} \alpha^* \omega \, \mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_{k-1} \\ &= \int_{\mathbb{R}^{k-1}} f_1(0, u_1, \ldots, u_{k-1}) \, \mathrm{d}u_1 \cdots \mathrm{d}u_{k-1} \,. \end{aligned}$$

En vue de (7.43) on peut conclure que (7.42) est établie. ♦

REMARQUE. La formule (7.42) est un cas très simple du théorème de Stokes généralisé ; d'autre part, (7.42) sera utilisée dans la démonstration de ce dernier.

7.8.8 Un théorème de convergence

Soit $\{\omega_n\}$, $n = 1, 2, \ldots$, une suite de k-formes ($k \geqslant 1$) dans une partie ouverte G de $\mathbb{R}^N$; si ω est une k-forme dans G, on dira que $\omega_n \to \omega$ (lorsque $n \to \infty$) dans une partie A de G si

$$\lim_{n\to\infty} f_{n,I}(x) = f_I(x) \,, \quad x \in A \,,$$

où les écritures canoniques pour ω_n et ω sont

$$\omega_n = \sum_I f_{n,I}\,\mathrm{d}x_I\,, \quad \omega = \sum_I f_I\,\mathrm{d}x_I\,.$$

Les notations étant fixées comme ci-dessus, énonçons le *théorème de convergence.*

THÉORÈME. *Soit X une variété différentielle orientée de dimension k dans G ; si $\omega_n \to \omega$ (lorsque $n \to \infty$) dans $A \subset X$ et si*

$$|f_{n,I}(x)| \leqslant M(x)\,, \quad x \in A \ (n \geqslant 1,\ I \in S(k;N))$$

où $\int_A M\,\mathrm{d}\sigma^X < \infty$, σ^X étant la mesure naturelle de X, alors ω_n, ω sont intégrables dans A et

$$\lim_{n\to\infty} \int_A \omega_n = \int_A \omega\,.$$

DÉMONSTRATION. Introduisons la notation

$$|\omega_n| = \left\{\sum_I |f_{n,I}|^2\right\}^{1/2}, \quad |\omega| = \left\{\sum_I |f_I|^2\right\}^{1/2};$$

puisque

$$|\omega_n| \leqslant CM\,, \quad |\omega| \leqslant CM\,, \quad C = \sqrt{\binom{N}{k}}\,,$$

on conclut que ω_n, ω sont intégrables dans A (§ 7.8.4) ; de (7.36) du paragraphe 7.8.4 on obtient que

$$\left|\int_A \omega_n - \int_A \omega\right| \leqslant \int_A |\omega_n - \omega|\,\mathrm{d}\sigma^X$$

où

$$|\omega_n - \omega| = \left\{\sum_I |f_{n,I} - f_I|^2\right\}^{1/2} \leqslant 2CM\,.$$

Le théorème de convergence dominée de Lebesgue donne que

$$\int_A |\omega_n - \omega|\,\mathrm{d}\sigma^X \longrightarrow 0\,, \quad n \to \infty$$

ce qui entraîne la conclusion désirée. ♦

REMARQUE. Un cas simple du théorème ci-dessus est le suivant : si $\sigma^X(A) < \infty$ et $|f_{n,I}(x)| \leqslant M$, $x \in A$ $(n \geqslant 1,\ I \in S(k;N))$, où M est une constante alors $\omega_n \to \omega$ $(n \to \infty)$ dans $A \subset X$ entraîne que $\int_A \omega_n \to \int_A \omega$ $(n \to \infty)$.

7.9 THÉORÈME DE STOKES GÉNÉRALISÉ

7.9.1 Proposition (*théorème de Stokes généralisé*)

Soit X une variété différentielle orientée de dimension $k \geqslant 2$ de classe C^1 contenue dans une partie ouverte G de $\mathbb{R}^N$. Soit Ω une partie de X qui est X-ouverte, bornée et telle que $\overline{\Omega}$, l'adhérence de Ω dans $\mathbb{R}^N$, est contenue dans X ; supposons que $\sigma(\partial\Omega) < \infty$ où σ est la mesure naturelle de la variété $\partial\Omega$ (bord régulier de Ω dans X, § 7.7.6) orientée selon l'orientation induite de X ; de plus, si $S = S_\Omega$ est la partie non régulière de $\mathrm{bord}(\Omega)$, *nous supposons que*

$$\lim_{\delta \to 0} \frac{1}{\delta} \sigma^X(S_\delta) = 0$$

où $S_\delta = \{x \in X : d(x, S) \leqslant \delta\}$ et σ^X est la mesure naturelle de X. Alors pour toute $(k-1)$-forme ω de classe C^1 dans G, on a

$$\int_\Omega \mathrm{d}\omega = \int_{\partial\Omega} \omega . \tag{7.44}$$

REMARQUES

Si Ω satisfait aux conditions de la proposition 7.9.1, nous dirons que Ω est une partie ***stokienne*** de X. En utilisant les méthodes du paragraphe 7.3.9. on peut démontrer facilement que si $S = S_\Omega$ est contenue dans une variété différentielle de classe C^1 de dimension $(k-2)$ dans X alors $(1/\delta)\,\sigma^X(S_\delta) \to 0$ $(\delta \to 0)$; ainsi, une condition suffisante pour qu'une partie X-ouverte $\Omega \in X$ soit une partie stokienne est que :

(i) Ω borné, $\overline{\Omega} \subset X$, $\sigma(\partial\Omega) < \infty$ $(\sigma = \sigma^{\partial\Omega})$;
(ii) $S = \mathrm{bord}(\Omega) \setminus \partial\Omega$ est contenu dans une réunion finie des sous-variétés différentielles de classe C^1 de dimension $(k-2)$.

Si X est une surface régulière orientée dans $\mathbb{R}^3$ et Ω est un domaine régulier dans X dans le sens du paragraphe 6.5.2 alors Ω est une partie stokienne de X ; ainsi, la proposition 6.5.4 est une conséquence de la proposition 7.9.1.

Notons un corollaire important de la proposition 7.9.1.

7.9.2 Corollaire

Les hypothèses sur ω étant comme dans la proposition 7.9.1, on a que

$$\int_X \mathrm{d}\omega = 0$$

pour toute variété différentielle compacte orientée X de dimension k de classe C^1 contenue dans G.

En effet, si X est compact, $\Omega = X$ est une partie X-ouverte bornée avec $\overline{\Omega} = \Omega$ et bord $\Omega = \partial\Omega = \emptyset$ ce qui donne le corollaire comme un cas particulier de (7.44). ♦

DÉMONSTRATION DE LA PROPOSITION 7.9.1

Nous faisons la démonstration sous *l'hypothèse* que la variété X est de *classe* C^2 ; nous indiquerons l'endroit exact où cette hypothèse est nécessaire pour notre démonstration ; dans les remarques du paragraphe 7.9.3 nous expliquons comment on peut établir la proposition sous l'hypothèse C^1.

Notre démonstration est similaire à celle donnée pour la proposition 7.4.6 (théorème de la divergence dans $\mathbb{R}^N$). Nous commençons par la remarque qu'il suffit de démontrer (7.44) pour une $(k-1)$-forme ω de classe C^1 définie partout dans $\mathbb{R}^N$ telle que $\operatorname{supp}\omega$ est compact et $X \cap \operatorname{supp}\omega = K$ est compact ; le raisonnement nécessaire pour justifier cette remarque est donné dans la démonstration mentionnée de la proposition 7.4.6. Ecrivons $S = S_\Omega$ pour la partie non régulière du bord(Ω) ; S est compact. Pour la première étape de la démonstration nous supposons que $K \cap S = \emptyset$ ou $K = X \cap \operatorname{supp}\omega$.

Pour chaque $x \in K$ nous choisissons un voisinage X-ouvert V_x de x, *disjoint de* S, tel que $V_x = \psi_x(U_x)$ où (ψ_x, U_x) est un paramétrage local positivement orienté (de classe C^2) ; ces choix sont faits comme suit :

(i) si $x \in K \cap \Omega$, $V_x \subset \Omega$;
(ii) si $x \in K \cap \operatorname{ext}_X \Omega$, $V_x \subset \operatorname{ext}_X \Omega$; ici $\operatorname{ext}_X \Omega = X \setminus \overline{\Omega}$;
(iii) si $x \in K \cap \partial\Omega$, on prend V_x, (ψ_x, U_x) tels que le paramétrage local (ψ_x, U_x) est adapté à $x \in \partial\Omega$ (§ 7.7.6).

Chaque V_x est de la forme $W_x \cap X$ où W_x est une partie ouverte de $\mathbb{R}^N$. Soit $\{\theta_i\}_{i\geqslant 1}$ une partition de l'unité subordonnée à $\{W_x\}_{x\in K}$. Puisque K est compact, il existe un entier $n \geqslant 1$ et une partie ouverte A avec $K \subset A$ tels que

$$\theta_1(x) + \cdots + \theta_n(x) = 1\,, \quad x \in A$$

(prop. 7.2.2 (ii)). Supposons que $\operatorname{supp}\theta_i \subset W_{x_i}$; alors

$$(\operatorname{supp}\theta_i) \cap X \subset W_{x_i} \cap X = V_{x_i}\,.$$

Vérifions que, pour $1 \leqslant i \leqslant n$,

$$\int_\Omega \mathrm{d}(\theta_i\omega) = \int_{\partial\Omega} \theta_i\omega\,. \tag{7.45}$$

En effet, si $\operatorname{supp}\theta_i \subset W_{x_i}$ où x_i est du type (i) ou (ii) ci-dessus, alors le second membre de (7.45) est nul car dans ces cas $\theta_i\omega = 0$ sur $\partial\Omega$; pour x_i du type (ii), $V_{x_i} \cap \Omega = \emptyset$ de sorte que le premier membre de (7.45) est alors également nul ; pour x_i du type (i), $V_{x_i} \subset \Omega$, $(\operatorname{supp}\theta_i\,\omega) \cap X \subset V_{x_i}$ et

$$\int_\Omega \mathrm{d}(\theta_i\,\omega) = \int_{V_{x_i}} \mathrm{d}(\theta_i\,\omega) = \int_{U_{x_i}} \psi_{x_i}^*(\mathrm{d}\theta_i\,\omega) = \int_{U_{x_i}} \mathrm{d}\widetilde{\omega}_i \tag{7.46}$$

où $\widetilde{\omega}_i = \psi_{x_i}^*(\theta_i\,\omega)$ est une $(k-1)$-forme (de classe C^1) à support compact dans U_{x_i} ; en écrivant

$$\widetilde{\omega}_i = \sum_{j=1}^{k} (-1)^{j-1} f_j \, \mathrm{d}u_1 \wedge \cdots \wedge \widehat{\mathrm{d}u_j} \wedge \cdots \wedge \mathrm{d}u_k$$

on aura

$$\mathrm{d}\widetilde{\omega}_i = \sum_{j=1}^{k} \mathrm{D}_j f_j \, \mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_k$$

avec

$$\int_{U_{x_i}} \mathrm{d}\widetilde{\omega}_i = \sum_{j=1}^{k} \int_{U_{x_i}} \mathrm{D}_j f_j \, \mathrm{d}u = 0$$

d'après la proposition 7.4.4 car les f_j sont de classe C^1 à support compact dans U_{x_i}. Ceci établit (7.45) pour x_i du type (i).

Si $\operatorname{supp}\theta_i \subset W_{x_i}$ où x_i est du type (iii), on fait le calcul suivant :

$$\begin{aligned} \int_{\Omega} \mathrm{d}(\theta_i\,\omega) &= \int_{V_{x_i}\cap\Omega} \mathrm{d}(\theta_i\,\omega) = \int_{U_{x_i}\cap H_k} \psi_{x_i}^*(\mathrm{d}\theta_i\,\omega) \\ &= \int_{U_{x_i}\cap H_k} \mathrm{d}\big(\psi_{x_i}^*(\theta_i\,\omega)\big) = \int_{H_k} \mathrm{d}\widetilde{\omega}_i \end{aligned} \tag{7.47}$$

où $H_k = \big\{(u_1, \ldots, u_k) : u_1 < 0\big\}$ et $\widetilde{\omega}_i$ est une $(k-1)$-forme (de classe C^1) à support compact obtenue en prolongeant $\psi_{x_i}^*(\theta_i\,\omega)$ en dehors de U_{x_i} en l'égalant à zéro sur $U_{x_i}^c$; puisque $\psi_{x_i}^*(\theta_i\,\omega)$ est à support compact contenu dans U_{x_i} ceci ne pose pas de problèmes. Mais, il faut noter ici que (7.46) et (7.47) sont obtenues grâce à la formule

$$\psi_{x_i}^*\big(\mathrm{d}(\theta_i\,\omega)\big) = \mathrm{d}\big(\psi_{x_i}^*(\theta_i\,\omega)\big)$$

valable lorsque ψ_{x_i} est de classe C^2 (§ 7.6.5(ii)) ; c'est seulement ici que la condition C^2 joue un rôle ; elle garantit aussi que $\psi_{x_i}^*(\theta_i\,\omega)$ est de classe C^1.

Pour calculer $\int_{\partial\Omega} \theta_i\,\omega$ nous reprenons les notations et les considérations des paragraphes 7.7.6 et 7.8.7 ; on a

$$\alpha(u_1, \ldots, u_{k-1}) = (0, u_1, \ldots, u_{k-1})$$

avec $\alpha : \mathbb{R}^{k-1} \to \partial H_k = \operatorname{Fr} H_k$; écrivons

$$\widehat{\psi}_{x_i} = \psi_{x_i} \circ \alpha\,, \quad \widehat{U}_{x_i} = \alpha^{-1}(U_{x_i} \cap \partial H_k)\,, \quad \widehat{\psi}_{x_i} : \widehat{U}_{x_i} \longrightarrow \widehat{V}_{x_i} = \widehat{\psi}_{x_i}(\widehat{U}_{x_i})\,.$$

$(\widehat{\psi}_{x_i}, \widehat{U}_{x_i})$ donne un paramétrage local de $\widehat{V}_{x_i} \subset \partial\Omega$ dans le sens de l'orientation induite de $\partial\Omega$. Alors,

$$\begin{aligned} \int_{\partial\Omega} \theta_i\,\omega &= \int_{\widehat{V}_{x_i}} \theta_i\,\omega = \int_{\widehat{U}_{x_i}} \widehat{\psi}_{x_i}^*(\theta_i\,\omega) = \int_{\widehat{U}_{x_i}} \alpha^*\big(\psi_{x_i}^*\theta_i\,\omega\big) \\ &= \int_{\mathbb{R}^{k-1}} \alpha^*(\widetilde{\omega}_i) = \int_{\partial H_k} \widetilde{\omega}_i\,. \end{aligned} \tag{7.48}$$

D'après (7.42) on a

$$\int_{H_k} \mathrm{d}\widetilde{\omega}_i = \int_{\partial H_k} \widetilde{\omega}_i$$

ce qui donne, au vu de (7.47) et (7.48), l'égalité (7.45) dans ce cas aussi.

Notons que

$$\sum_{i=1}^{n} \theta_i \, \omega = \omega$$

dans X car si $x \in A$, $\sum_{i=1}^{n} \theta_i(x) = 1$ et si $x \in X \cap A^c$, $\omega(x) = 0$, étant donné que $(\operatorname{supp} \omega) \cap X = K \subset A$.

En faisant la somme sur $i = 1, \ldots, n$ dans (7.45) on obtient (7.44) dans le cas où $K \cap S = \emptyset$.

Supposons maintenant que $K \cap S \neq \emptyset$. Pour chaque $\delta > 0$, on choisit $\varphi_\delta \in C^\infty(\mathbb{R}^N)$ comme dans la preuve de la proposition 7.4.6 ; on aura

$$\operatorname{supp}(\varphi_\delta \, \omega) \subset K \setminus S \,,$$

d'où, en appliquant (7.44) à $\varphi_\delta \, \omega$, on a

$$\int_\Omega \mathrm{d}(\varphi_\delta \, \omega) = \int_{\partial\Omega} \varphi_\delta \, \omega \,. \tag{7.49}$$

Démontrons que

$$\lim_{\delta \to 0} \int_{\partial\Omega} \varphi_\delta \, \omega = \int_{\partial\Omega} \omega \,. \tag{7.50}$$

En effet, (7.50) est une conséquence du théorème de convergence (§ 7.8.8) ; on a $\sigma(\partial\Omega) < \infty$ par hypothèse et si

$$\omega = \sum_{I \in S(k-1;N)} f_I \, \mathrm{d}x_I$$

on aura

$$\varphi_\delta \, \omega = \sum_I \varphi_\delta f_I \, \mathrm{d}x_I$$

d'où l'existence d'un M', $0 < M' < \infty$, tel que

$$\left|\varphi_\delta f_I(x)\right| \leqslant \left|f_I(x)\right| \leqslant M' \,, \quad x \in \overline{\Omega} \,,\ I \in S(k-1;N) \,,$$

car les f_I sont de classe C^1 et $\overline{\Omega}$ est compact. Puisque

$$\lim_{\delta \to 0} \varphi_\delta(x) = \begin{cases} 0 & \text{si } x \in S \\ 1 & \text{si } x \in S^c \end{cases}$$

on a $\lim_{\delta \to 0} \varphi_\delta \, \omega = \omega$ sur $\partial\Omega$ (car $\partial\Omega \subset S^c$).

D'autre part,

$$\int_\Omega \mathrm{d}(\varphi_\delta \, \omega) = \int_\Omega \mathrm{d}\varphi_\delta \wedge \omega + \int_\Omega \varphi_\delta \wedge \mathrm{d}\omega \,. \tag{7.51}$$

On démontre que

$$\lim_{\delta\to 0}\int_\Omega \varphi_\delta\wedge\omega=\int_\Omega \mathrm{d}\omega\,, \tag{7.52}$$

exactement comme on l'a fait pour (7.50).

Démontrons finalement que

$$\lim_{\delta\to 0}\int_\Omega \mathrm{d}\varphi_\delta\wedge\omega=0\,. \tag{7.53}$$

Ici on utilise l'inégalité (7.37) et le raisonnement à la fin de la démonstration de la proposition 7.4.6 ; si

$$\omega=\sum_{I\in S(k-1;N)} f_I\,\mathrm{d}x_I$$

on a

$$\mathrm{d}\varphi_\delta\wedge\omega=\sum_{j=1}^{N}\mathrm{D}_j\varphi_\delta\,\mathrm{d}x_j\wedge\omega=\sum_{I,j}\mathrm{D}_j\varphi_\delta\cdot f_I\,\mathrm{d}x_j\wedge\mathrm{d}x_I\,.$$

Puisque $\left|\mathrm{D}_j\varphi_\delta(x)\right|\leqslant 2M/\delta$, $x\in\mathbb{R}^N$, on obtient de (7.37) que, pour un $C>0$ approprié, on aura

$$\begin{aligned}\left|\int_\Omega \mathrm{d}\varphi_\delta\wedge\omega\right|&=\left|\int_{\Omega\cap\{x:\delta/2\leqslant \mathrm{d}(x,S)\leqslant\delta\}}\mathrm{d}\varphi_\delta\wedge\omega\right|\\&\leqslant\frac{C}{\delta}\,\sigma^X\{x:\mathrm{d}(x,S)\leqslant\delta\}\longrightarrow 0\text{ lorsque }\delta\to 0\end{aligned}$$

selon notre hypothèse sur S. Ceci établit (7.53).

Au vu de (7.49), (7.50), (7.51) et (7.52), la preuve est achevée. ♦

7.9.3 Remarques

Pour éliminer l'hypothèse que X est de classe C^2 on peut suivre une méthode élégante (due à E. Cartan et présentée dans Whitney [43, pp. 103-108]). Elle consiste à travailler avec les p-formes dites ***régulières*** ; on dit que ω est une p-forme régulière dans une partie ouverte G de $\mathbb{R}^N$ si elle est continue et s'il existe une $(p+1)$-forme continue ω' telle que pour tout parallélépipède P de dimension $(p+1)$ dans G on a

$$\int_{\partial P}\omega=\int_P\omega'\,.$$

On peut démontrer l'unicité de ω' (si elle existe, c'est-à-dire si ω est régulière) et on pose $\mathrm{d}\omega=\omega'$; si ω est de classe C^1 alors ω est régulière et la forme ω' coïncide avec $\mathrm{d}\omega$ de la définition initiale. Les formules usuelles concernant $\mathrm{d}\omega$ se généralisent et on obtient en particulier

$$\mathrm{d}(\mathrm{d}\omega)=0\,,\quad \mathrm{d}(\psi^*\omega)=\psi^*(\mathrm{d}\omega)$$

sous la seule hypothèse que ω soit régulière et ψ de classe C^1. Ainsi, avec cette définition de $\mathrm{d}\omega$, (7.47) est justifié et la démonstration présentée pour la proposition 7.9.1 reste valable pour X de classe C^1.

Il faut remarquer que dans la vaste majorité des applications les variétés X sont même de classe C^∞.

Au vu du paragraphe 7.8.5 où nous avons identifié l'intégration d'une $(N-1)$-forme sur une hypersurface orientée comme le flux d'un champ vectoriel (cf. (7.38)), la proposition 7.9.1 généralise le théorème de la divergence dans $\mathbb{R}^N$, ce dernier étant le cas $k = (N-1)$ du théorème de Stokes généralisé. Ainsi, la proposition 7.9.1 contient l'ensemble des trois théorèmes associés avec les noms de Green, Gauss et Stokes que l'on a vu dans le chapitre 6.

La proposition 7.9.1 est valable aussi pour $k = 1$ pourvu que l'on considère $\partial\Omega$ comme un ensemble discret (variété de dimension 0) dans X de dimension 1 considérée comme une courbe régulière simple (si X est connexe) dans $\mathbb{R}^N$; alors (7.44) est un cas particulier d'une formule connue de la théorie des intégrales curvilignes :

$$\int_\gamma \nabla u(x) \cdot \mathrm{d}\vec{s} = u\big(\gamma(1)\big) - u\big(\gamma(0)\big)$$

où $\gamma : [0\,,\,1] \to \mathbb{R}^N$ est un chemin de classe C^1 et u est une fonction réelle de classe C^1. Notons que ω s'identifie avec u (cas $k = 0$) et

$$\mathrm{d}\omega = \mathrm{D}_1 u\,\mathrm{d}x_1 + \cdots + \mathrm{D}_N u\,\mathrm{d}x_N \ ;$$

se rappeler que

$$\int_\gamma (\mathrm{D}_1 u\,\mathrm{d}x_1 + \cdots + \mathrm{D}_N u\,\mathrm{d}x_N) = \int_\gamma \nabla u(x) \cdot \mathrm{d}\vec{s}.$$

Nous donnons maintenant brièvement deux autres formulations du théorème de Stokes souvent utilisées ; nous n'entrons pas dans les détails de ces formulations, l'objectif étant seulement d'informer le lecteur de l'existence de ces variantes.

7.9.4 Version simpliciale du théorème de Stokes

Cette version est presque entièrement algébrique. L'élément analytique principal au-delà des maniements des formes différentielles est la définition suivante :

$$\int_\psi \omega = \int_D \psi^* \omega$$

où ω est une k-forme différentielle continue dans une partie ouverte G de $\mathbb{R}^N$ et $\psi : \overline{D} \to G$ est une application de classe C^1 définie dans l'adhérence $\overline{D}$ d'une partie ouverte bornée D de $\mathbb{R}^k$. Puisque $\psi^* \omega$ est de la forme $g\,\mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_k$ avec $g \in C(\overline{D})$, l'intégrale concernée existe et on a

$$\int_\psi \omega = \int_D g(u)\,\mathrm{d}u\,.$$

Dans la suite de ce paragraphe, $G \subset \mathbb{R}^N$ sera toujours une partie ouverte et les applications ψ seront toujours de classe C^1 au moins.

Soient $e_0 = 0$, $e_1 = (1, 0, \ldots, 0)$, $\ldots$, $e_k = (0, 0, \ldots, 1)$, $\{e_1, \ldots, e_k\}$ étant la base canonique de $\mathbb{R}^k$. Soit Q_k le ***k-simplexe standard*** de $\mathbb{R}^k$:

$$Q_k = \left\{ \sum_{i=1}^{k} u_i e_i : u_i \geqslant 0\,,\ i = 1, \ldots, k\,,\ \sum_{i=1}^{k} u_i \leqslant 1 \right\}.$$

Si $a_0, a_1, \ldots, a_k$ sont $(k+1)$-vecteurs dans $\mathbb{R}^m$, notons par $\sigma = [a_0, \ldots, a_k]$ l'application $\sigma : Q_k \to \mathbb{R}^m$ définie comme suit :

$$\sigma(u) = a_0 + Tu\,, \quad u \in Q_k\,,$$

où T est l'unique application linéaire de $\mathbb{R}^k$ dans $\mathbb{R}^m$ définie par $Te_i = a_i - a_0$.

Un ***k-simplexe singulier*** dans G est une application $\psi : Q_k \to G$ de la forme $\psi = \varphi \circ \sigma$ où $\varphi : \sigma(Q_k) \to G$ et $\sigma = [a_0, \ldots, a_k] : Q_k \to \mathbb{R}^m$. Une ***k-chaîne singulière*** ψ dans G est une combinaison linéaire formelle des k-simplexes singuliers dans G :

$$\psi = \sum_{i=1}^{n} c_i \psi_i$$

avec $c_i \in \mathbb{R}$, ψ_i un k-simplexe singulier dans G, $i = 1, \ldots, n$; on pose alors

$$\int_{\psi} \omega = \sum_{i=1}^{n} c_i \int_{\psi_i} \omega\,,$$

ω étant une k-forme continue dans G. Pour un k-simplexe singulier $\psi : Q_k \to G$, on définit $\partial\psi$ comme une $(k-1)$-chaîne singulière dans G telle que

$$\int_{\partial\psi} \omega = \int_{\partial \mathring{Q}_k} \psi^* \omega$$

pour toute $(k-1)$-forme ω continue dans G, $\partial \mathring{Q}_k$ étant orienté par l'orientation induite de l'orientation canonique de $\mathring{Q}_k \subset \mathbb{R}^k$. Pour cela, il suffit de poser

$$\partial\psi = \sum_{i=0}^{k} (-1)^i \psi_i$$

avec $\psi_i = \varphi \circ \sigma_i$, $\sigma_i = [a_0, \ldots, \widehat{a_i}, \ldots, a_k]$ (suppression de a_i) où $\psi = \varphi \circ \sigma$, $\sigma = [a_0, \ldots, a_k]$. Nous ne présentons pas la vérification directe donnant la propriété indiquée de $\int_{\partial\psi} \omega$. On définit $\partial\psi$ pour une k-chaîne singulière par linéarité.

Le théorème de Stokes maintenant a la forme suivante. *Si ω est une $(k-1)$-forme continue et ψ est une k-chaîne singulière dans G alors*

$$\int_{\psi} \mathrm{d}\omega = \int_{\partial\psi} \omega .$$

Pour la preuve, il suffit de considérer le cas d'un k-simplexe singulier $\psi : Q_k \to G$ (de classe C^2) et noter les égalités suivantes :

$$\int_{\psi} \mathrm{d}\omega \overset{1}{=} \int_{\dot{Q}_k} \psi^*(\mathrm{d}\omega) \overset{2}{=} \int_{\dot{Q}_k} \mathrm{d}(\psi^*\,\omega) \overset{3}{=} \int_{\partial\dot{Q}_k} \psi^*\,\omega \overset{4}{=} \int_{\partial\psi} \omega .$$

Les égalités 1 et 4 résultent des définitions ; l'égalité 2 est une propriété générale des formes différentielles ; finalement, l'égalité 3 est le théorème de la divergence formulé en termes de formes différentielles et appliqué à l'ensemble ouvert particulier $\dot{Q}_k \subset \mathbb{R}^k$ (§ 7.9.3).

La version simpliciale du théorème de Stokes est surtout utilisée dans la topologie algébrique. Il y a plusieurs variations possibles à cette version ; on peut remplacer les simplexes par des cubes ou par d'autres polyèdres. Un exposé classique est celui donné dans de Rham [34].

7.9.5 Théorème de Stokes pour les variétés à bord

La notion d'une variété à bord est comprise dans la définition de l'ensemble $\overline{\Omega}$ (contenu dans X) de la proposition 7.9.1 ; on dit que $\overline{\Omega}$ est une variété de dimension k avec bord $\partial\Omega$ (en supposant que S_Ω, la partie non régulière de bord(Ω), est vide). Précisons cette notion.

Soit E une partie de $\mathbb{R}^N$; on dit que E est une ***variété différentielle à bord*** de dimension $k \geqslant 1$ de classe C^1 dans $\mathbb{R}^N$ si les conditions suivantes sont vérifiées : chaque point $x \in E$ possède un voisinage E-ouvert V qui peut être paramétré par (ψ, U) où U est une partie ouverte soit de $\mathbb{R}^k$ soit de $\overline{H}_k = \{(x_1, \ldots, x_k) : x_1 \leqslant 0\}$ et $\psi : U \to V$ satisfait aux conditions (habituelles) suivantes :

- ψ est un homéomorphisme entre U et V,
- ψ est de classe C^1,
- $\operatorname{rang} \psi'(u) = k$, $u \in U$.

Si x est un point de E qui admet un voisinage E-ouvert paramétré par (ψ, U) où U et une partie ouverte de $\mathbb{R}^k$ on dit que x est un ***point intérieur*** de E ; sinon on dit que x est un point appartenant au ***bord*** de E noté ∂E. Ainsi,

$$E = (\operatorname{Int} E) \cup \partial E ,$$

$\operatorname{Int} E$ étant l'ensemble formé par les points intérieurs de E. On démontre que $\operatorname{Int} E$, ∂E sont disjoints et que ∂E est une variété différentielle (sans bord) de dimension $(k-1)$ de classe C^1. Si $\partial E = \emptyset$, E est une variété différentielle

(dite ***sans bord***) du type que nous avons considéré jusqu'à maintenant. Noter que $\operatorname{Int} E$ est toujours non vide et forme une variété différentielle sans bord de dimension k de classe C^1 ; en fait E est l'adhérence dans $\mathbb{R}^N$ de $\operatorname{Int} E$:

$$E = \overline{\operatorname{Int} E}, \quad \partial E = E \setminus \operatorname{Int} E$$

pourvu que $\partial E \neq \emptyset$. Si $\partial E = \emptyset$ et E est une variété différentielle sans bord de dimension k dans $\mathbb{R}^k$ alors $E = \operatorname{Int} E$ mais $\overline{E}$ n'est pas nécessairement égal à E ; évidemment, ce sera le cas si E est fermé (par exemple, compact) en tant que sous-ensemble de $\mathbb{R}^N$.

Un exemple simple illustre bien les définitions données. Soit $E = \overline{B}(0;1)$, la boule unité fermée centrée à 0, dans $\mathbb{R}^N$; alors

$$\operatorname{Int} E = B(0;1) = \mathring{E}, \quad \partial E = S(0;1) = \operatorname{Fr} E.$$

En partant d'une variété différentielle X (sans bord) de dimension k de classe C^1 dans $\mathbb{R}^N$, l'ensemble $\overline{X}$ peut être très compliqué ; en particulier, $\overline{X} \setminus X$ n'est pas nécessairement une variété différentielle. Néanmoins, on peut isoler une partie ∂X de $\overline{X} \setminus X$ (***bord régulier*** de X) qui sera une variété différentielle de dimension $(k-1)$ de classe C^1 et $E = X \cup \partial X$ sera une variété à bord avec $\partial E = \partial X$.

On peut introduire maintenant la notion d'une variété à bord E orientée et introduire l'orientation induite de ∂E. Les discussions sont tout à fait analogues à celles que nous avons données dans les paragraphes 7.7.3 et 7.7.6. La théorie de l'intégration et de la dérivation extérieure des formes se présente comme dans notre développement. Une version simple du théorème de Stokes dans ce cadre est la suivante : on a

$$\int_E \mathrm{d}\omega = \int_{\partial E} \omega$$

si ω est une $(k-1)$-forme de classe C^1 à support compact dans une partie ouverte G de $\mathbb{R}^N$ et E est une variété différentielle orientée de dimension k de classe C^1 dans G ayant le bord ∂E orientée par l'orientation induite. Dans cette formulation, il est difficile de permettre un « bord anguleux » car le bord ∂E est « lisse » par définition même d'une variété à bord. Par contre, on peut introduire une formulation assez générale en considérant $\partial X \subset \overline{X} \setminus X \subset G$ où X est une variété différentielle de dimension k (sans bord) dans G et ∂X est son bord régulier ; nous ne développerons pas cette formulation ici. Il est possible que l'on ait un théorème dans cette formulation ayant le même ordre de généralité que la proposition 7.9.1 ; par contre, la théorie générale des variétés à bord anguleux semble être assez compliquée.

7.9.6 Exercices

[1] Soient $f_i : \mathbb{R}^N \to \mathbb{R}$, $i = 1, \ldots, N$, des fonctions de classe C^1 paires : $f_i(-x) = f_i(x)$, $i \geqslant 1$, $x \in \mathbb{R}^N$. Posons

$$\omega = \sum_{i=1}^{N} (-1)^{i-1} f_i \, \mathrm{d}x_1 \wedge \cdots \wedge \widehat{\mathrm{d}x_i} \wedge \cdots \wedge \mathrm{d}x_N .$$

Démontrer que $\int_{S^{N-1}} \omega = 0$ où S^{N-1} est la sphère unité centrée en 0 et orientée par sa normale unitaire dirigée vers l'extérieur.
(Calculer $\mathrm{d}\omega = g \, \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_N$ et observer que $g(-x) = -g(x)$.)

[2] Soient ω, Ω, G comme dans la proposition 7.9.1 ; démontrer que si $f : G \to \mathbb{R}$ est une fonction de classe C^1 alors

$$\int_{\Omega} \mathrm{d}f \wedge \omega = \int_{\partial\Omega} f\omega - \int_{\Omega} f \, \mathrm{d}\omega .$$

[3] Soient α une p-forme, β une q-forme de classe C^1 dans une partie ouverte G de $\mathbb{R}^N$. Si Ω est une partie stokienne d'une variété différentielle orientée X de dimension $(p+q+1)$ de classe C^1 contenue dans G, démontrer que

$$\int_{\Omega} \mathrm{d}\alpha \wedge \beta = \int_{\partial\Omega} \alpha \wedge \beta + (-1)^{p+1} \int_{\Omega} \alpha \wedge \mathrm{d}\beta .$$

[4] Soit f une fonction harmonique dans $\mathbb{R}^N \setminus \{0\}$; posons

$$\omega_f = \sum_{i=1}^{N} (-1)^{i-1} \mathrm{D}_i f \, \mathrm{d}x_1 \wedge \cdots \wedge \widehat{\mathrm{d}x_i} \wedge \cdots \wedge \mathrm{d}x_N .$$

Vérifier que $\mathrm{d}\omega_f = 0$; si $f(x) = \|x\|^{-N+2}$, $x \neq 0$, f est harmonique dans $\mathbb{R}^N \setminus \{0\}$ et ω_f est une $(N-1)$-forme fermée qui n'est pas exacte car $\int_{S_{N-1}} \omega \neq 0$ où S_{N-1} est la sphère unitaire centrée en 0 orientée par sa normale unitaire dirigée vers l'extérieur.
(Pour calculer $\int_{S_{N-1}} \omega$ identifier l'intégrale comme $\int_{S_{N-1}} F \cdot \nu \, \mathrm{d}\sigma$ selon le paragraphe 7.8.5.)

[5] Soient G une partie ouverte de $\mathbb{R}^N$, ω une k-forme continue dans G, $\psi = \varphi \circ \sigma$, $\sigma = [a_0, \ldots, a_k]$, un k-simplexe singulier dans G (§ 7.9.4). Démontrer que $\int_{\psi} \omega = 0$ si $a_1 - a_0, \ldots, a_k - a_0$ sont linéairement dépendants. Soient $\pi\sigma = [a_{\pi(0)}, \ldots, a_{\pi(k)}]$ où π est une permutation de $\{0, 1, \ldots, k\}$; démontrer que si $\pi\psi = \varphi \circ \pi\sigma$ alors

$$\int_{\pi\psi} \omega = \varepsilon(\pi) \int_{\psi} \omega$$

où $\varepsilon(\pi)$ est la signature de π.

7.10 FORMES FERMÉES, FORMES EXACTES : RETOUR

7.10.1 Introduction

Soit G une partie ouverte non vide dans $\mathbb{R}^N$. Dans le paragraphe 7.1.4 on a défini les notions de formes fermées et de formes exactes dans G ; une k-forme ($k \geqslant 0$) ω dans G est fermée si elle est de classe C^1 et si $\mathrm{d}\omega = 0$; ω est exacte s'il existe une $(k-1)$-forme π (une primitive de ω) de classe C^1 telle que $\mathrm{d}\pi = \omega$ (ici $k \geqslant 1$). Si une k-forme ω est exacte et si ω est de classe C^1 alors sa primitive π sera de classe C^2 et on aura $\mathrm{d}(\mathrm{d}\pi) = \mathrm{d}\omega = 0$; donc une telle forme exacte ω est fermée. Dans cette section nous étudions la question réciproque : quand peut-on affirmer qu'une forme fermée est exacte ? Cette question s'est présentée plusieurs fois dans le contexte de l'analyse vectorielle au chapitre 6 et elle a été discutée brièvement dans le paragraphe 7.1.4.

Nous commençons par l'introduction d'une notion importante.

7.10.2 Homotopie des applications

Soient U et V deux parties ouvertes de $\mathbb{R}^m$ et $\mathbb{R}^n$ respectivement et a, b deux applications de classe C^r de U dans V. On dit que a, b sont C^r-***homotopes*** (simplement ***homotopes*** si $r = 0$) s'il existe une application H : $U \times [0, 1] \to V$ de classe C^r telle que

$$H(x,0) = a(x), \quad H(x,1) = b(x), \quad x \in U.$$

H s'appelle une ***homotopie***. Noter que $U \times [0, 1]$ n'est pas ouvert dans $\mathbb{R}^{m+1}$; donc pour $r > 0$ la condition de classe C^r doit être interprétée de manière habituelle : il existe $\widetilde{H} \in C^r(W, \mathbb{R}^n)$ avec $W \supset U \times [0, 1]$, W ouvert dans $\mathbb{R}^{m+1}$, tel que $\widetilde{H}|_{U \times [0,1]} = H$; nous confondons H et $\widetilde{H}$.

La notion d'homotopie est fondamentale ; dans le présent contexte, elle précise l'idée de la « déformation continue » de l'application a pour aboutir a l'application b.

Nous établissons un lemme qui nous sera utile dans la suite. En retenant les notations précédentes avec $r = 2$, *nous supposons que a, b sont C^2-homotopes.*

LEMME. *A chaque k-forme continue ω dans V, $k \geqslant 1$, on peut faire correspondre une $(k-1)$-forme continue $\mathcal{P}\omega$ dans U telle que $\omega \mapsto \mathcal{P}\omega$ est linéaire et*

$$\mathrm{d}(\mathcal{P}\omega) + \mathcal{P}(\mathrm{d}\omega) = b^*\omega - a^*\omega \tag{7.54}$$

si ω est de classe C^1 dans V ($\mathcal{P}\omega$ est alors de classe C^1 dans U).

DÉMONSTRATION. Nous considérons d'abord une situation très particulière : posons

$$\alpha(x) = (x, 0)\,, \quad \beta(x) = (x, 1)\,, \quad x \in U \subset \mathbb{R}^m\,.$$

Nous considérons α, β comme des applications de classe C^∞ de U dans $W \supset U \times [0\,,\,1]$. Il est clair que α, β sont C^∞-homotopes ; il suffit de prendre comme homotopie $h(x,t) = (x,t)$, $h : U \times [0\,,\,1] \to W \subset \mathbb{R}^{m+1}$. Nous prenons les fonctions coordonnées de W comme $x_1, \ldots, x_m, t$ et celles dans U comme $x_1, \ldots, x_m$. Une k-forme ω ($k \geqslant 1$) dans W a l'écriture canonique

$$\omega = \sum_{I \in S(k;m)} f_I \, \mathrm{d}x_I + \sum_{J \in S(k-1;m)} g_J \, \mathrm{d}x_J \wedge \mathrm{d}t$$

où f_I et g_J sont des fonctions de $x = (x_1, \ldots, x_m)$ et t. Posons

$$P\omega = \sum_{J \in S(k-1;m)} (-1)^{k-1} (\mathcal{J} g_J) \, \mathrm{d}x_J$$

où

$$(\mathcal{J} g_J)(x) = \int_0^1 g_J(x,t) \, \mathrm{d}t\,.$$

Il est clair que $P\omega$ donne (par restriction) une $(k-1)$-forme dans U pour chaque k-forme ω dans W et la correspondance $\omega \mapsto P\omega$ est linéaire ; aussi, si ω est de classe C^r, $P\omega$ sera de classe C^r.

Nous vérifions maintenant par un calcul direct que l'analogue de (7.54) est vérifié pour P :

$$\mathrm{d}(P\omega) + P(\mathrm{d}\omega) = \beta^*\omega - \alpha^*\omega\,. \tag{7.55}$$

Par exemple, si $\omega = f$ est une 0-forme de classe C^1 dans W on a

$$\mathrm{d}f = \sum_{i=1}^m \mathrm{D}_i f \, \mathrm{d}x_i + \frac{\partial f}{\partial t} \, \mathrm{d}t$$

$$P(\mathrm{d}f) = \mathcal{J}\left(\frac{\partial f}{\partial t}\right) = f \circ \beta - f \circ \alpha = \beta^* f - \alpha^* f\,.$$

Si ω est une k-forme de classe C^1 dans W du type $\omega = f \, \mathrm{d}x_I$ on aura $P\omega = 0$ et

$$\begin{aligned}
\mathrm{d}(P\omega) + P(\mathrm{d}\omega) = P(\mathrm{d}f \wedge \mathrm{d}x_I) &= P\left(\sum_{i=1}^m \mathrm{D}_i f \, \mathrm{d}x_i \wedge \mathrm{d}x_I + \frac{\partial f}{\partial t} \, \mathrm{d}t \wedge \mathrm{d}x_I\right) \\
&= (-1)^k P\left(\frac{\partial f}{\partial t} \, \mathrm{d}x_I \wedge \mathrm{d}t\right) = \mathcal{J}\left(\frac{\partial f}{\partial t}\right) \mathrm{d}x_I \\
&= (f \circ \beta - f \circ \alpha) \, \mathrm{d}x_I = \beta^* f - \alpha^* f\,.
\end{aligned}$$

Notons que $\alpha^*(\mathrm{d}x_i) = \mathrm{d}x_i$, $\alpha^*(\mathrm{d}t) = 0$, $i = 1, \ldots, m$, et de même pour β.

Si $\omega = g\,\mathrm{d}x_J \wedge \mathrm{d}t$ est une k-forme de classe C^1 dans W (donc J contient $(k-1)$ indices) on aura

$$\begin{aligned}\mathrm{d}(P\omega) + P(\mathrm{d}\omega) &= \mathrm{d}\Big((-1)^{k-1}\mathcal{J}(g)\,\mathrm{d}x_J\Big) + P\left(\sum_{i=1}^{m} \mathrm{D}_i g\,\mathrm{d}x_i \wedge \mathrm{d}x_J \wedge \mathrm{d}t\right)\\ &= (-1)^{k-1}\sum_{i=1}^{m} \mathrm{D}_i\big(\mathcal{J}(g)\big)\,\mathrm{d}x_i \wedge \mathrm{d}x_J + \sum_{i=1}^{m}(-1)^k \mathcal{J}(\mathrm{D}_i g)\,\mathrm{d}x_i \wedge \mathrm{d}x_J\\ &= 0 = \beta^*\omega - \alpha^*\omega\end{aligned}$$

car $\mathrm{D}_i\big(\mathcal{J}(g)\big) = \mathcal{J}(\mathrm{D}_i g)$ (échange de $\int_0^1$ et D_i) et $\beta^*\omega = 0 = \alpha^*\omega$ dans ce cas. Donc (7.55) est vérifiée pour toute k-forme ω de classe C^1 dans W.

Posons maintenant

$$\mathcal{P}\omega = P(H^*\omega)$$

en se rappelant que $H : W \to V$, $W \supset U \times [0\,,\,1]$; donc une k-forme ω dans V donne une $(k-1)$-forme $\mathcal{P}\omega$ dans U. La vérification de (7.54) est un calcul direct :

$$\begin{aligned}\mathrm{d}(\mathcal{P}\omega) + \mathcal{P}(\mathrm{d}\omega) &= \mathrm{d}(PH^*\omega) + PH^*(\mathrm{d}\omega)\\ &= \mathrm{d}(PH^*\omega) + P(\mathrm{d}H^*\omega)\\ &= \beta^*H^*\omega - \alpha^*H^*\omega \qquad \text{(d'après (7.55))}\\ &= \big(H\circ\beta\big)^*\omega - \big(H\circ\alpha\big)^*\omega = b^*\omega - a^*\omega\,. \qquad \blacklozenge\end{aligned}$$

7.10.3 Théorème de Poincaré

Retenons les notations du paragraphe 7.10.2. Une conséquence immédiate de (7.54) est la suivante : supposons que $U = V$ est une partie ouverte de $\mathbb{R}^N$ et $a(x) = p \in U$, $x \in U$ et $b(x) = x$, $x \in U$; autrement dit, a est une application constante de U dans U et b est l'application identité de U dans U ; si a, b sont C^2-homotopes (on dit alors que U est ***contractile***) alors toute k-forme fermée ($k \geqslant 1$) ω de classe C^1 dans U est exacte. En effet, $a^*\omega = 0$, $b^*\omega = \omega$; si $\mathrm{d}\omega = 0$, on aura de (7.54) l'existence d'une $(k-1)$-forme $\pi = \mathcal{P}\omega$ telle que $\mathrm{d}\pi = \omega$. Nous formulons cette conclusion explicitement. *Soit U une partie ouverte non vide contractile dans $\mathbb{R}^N$; alors toute k-forme fermée de classe C^1 dans U est exacte.*

Comme conséquence de cet énoncé, nous obtenons la proposition suivante.

7.10.4 Proposition (*théorème de Poincaré*)

Soit U une partie ouverte non vide étoilée dans $\mathbb{R}^N$; alors toute k-forme fermée de classe C^1 dans U est exacte.

DÉMONSTRATION. Soit p un point de U tel que U est étoilé par rapport à p ; posons

$$H(x,t) = tx + (1-t)p, \quad 0 \leqslant t \leqslant 1\,,\ x \in U\,.$$

Puisque $[p\,,\,x] \subset U$ pour tout $x \in U$, H est une application de $U \times [0\,,\,1] \to U$; noter que pour chaque $x \in U$, H est défini par la même formule pour $t \in\ ]-\varepsilon\,,\,1+\varepsilon[\,,\ \varepsilon = \varepsilon(x) > 0$, de sorte que H est bien de classe C^∞ dans $U \times [0\,,\,1]$. De plus,

$$H(x,0) = a(x) = p, \quad H(x,1) = b(x) = x\,, \quad x \in U\ ;$$

on voit donc que U est contractile. La proposition résulte de l'observation du paragraphe 7.10.3. ♦

7.10.5 Remarques

Comme indiqué dans le paragraphe 7.1.4, les questions concernant l'existence d'un potentiel scalaire ou d'un potentiel vecteur pour un champ vectoriel dans $\mathbb{R}^3$ sont des cas particuliers du problème général traité dans cette section. Notamment, la proposition 6.6.3 est un cas particulier du théorème de Poincaré (prop. 7.10.4).

Une formulation élégante du théorème de Poincaré est celle qui utilise la ***cohomologie de de Rham***. Nous pouvons l'expliquer brièvement ici, car nous avons maintenant toute la terminologie nécessaire. Soit U une partie ouverte non vide dans $\mathbb{R}^N$. Posons $\Omega^k(U)$ pour l'espace vectoriel réel formé de toutes les k-formes de classe C^∞ dans U, $k = 0,\ 1,\ \ldots,\ N$; on a alors une suite d'homomorphismes (applications linéaires) :

$$0 \longrightarrow \mathbb{R} \longrightarrow \Omega^0(U) \xrightarrow{\mathrm{d}} \Omega^1(U) \longrightarrow \ldots \xrightarrow{\mathrm{d}} \Omega^{N-1}(U) \xrightarrow{\mathrm{d}} \Omega^N(U) \longrightarrow 0 \tag{7.56}$$

Les homomorphismes $0 \to \mathbb{R}$, $\mathbb{R} \to \Omega^0(U)$ sont les injections banales : 0 va sur 0, la constante c va sur la fonction constante c ; de même, $\Omega^N(U) \to 0$ est l'homomorphisme banal qui envoie toute N-forme à 0. Le fait que $\mathrm{d}(\mathrm{d}\omega) = 0$ a comme conséquence la relation suivante : posons

$$Z^k(U) = \mathrm{Ker}\big(\Omega^k(U) \xrightarrow{\mathrm{d}} \Omega^{k+1}(U)\big)\,, \quad k \geqslant 0$$

$$B^k(U) = \mathrm{Im}\big(\Omega^{k-1}(U) \xrightarrow{\mathrm{d}} \Omega^k(U)\big)\,, \quad k \geqslant 1$$

alors $B^k(U)$ *est un sous-espace vectoriel de* $Z^k(U)$. Plus explicitement : $\omega \in Z^k(U)$ si $\omega \in \Omega^k(U)$ et $\mathrm{d}\omega = 0$; $\omega \in B^k(U)$ si $\omega \in \Omega^k(U)$ et il existe $\pi \in \Omega^{k-1}(U)$ tel que $\mathrm{d}\pi = 0$; donc $\omega \in Z^k(U)$ *si* ω *est une* k*-forme,* $k \geqslant 0$, *fermée* (de classe C^∞) et $\omega \in B^k(U)$ *si* ω *est une* k*-forme,* $k \geqslant 1$, *exacte*. On définit

$$H^k(U) = Z^k(U)/B^k(U)\,, \quad 1 \leqslant k \leqslant N\ ;$$

c'est l'espace vectoriel quotient de $Z^k(U)$ par le sous-espace $B^k(U)$; on définit

$$H^0(U) = \mathrm{Ker}\big(\Omega^0(U) \xrightarrow{\mathrm{d}} \Omega^1(U)\big) ;$$

$\omega \in H^0(U)$ est une fonction de classe C^∞ dans U telle que $\mathrm{d}\omega = 0$. Les espaces $H^k(U)$ sont par définition égaux à $\{0\}$ pour $k > N$ si $U \subset \mathbb{R}^N$; ceci correspond au fait que $\Omega^k(U) = \{0\}$ si $k > N$. Les espaces $H^k(U)$, $k \geqslant 0$, s'appellent les ***groupes de cohomologie de de Rham*** de U, même si chaque $H^k(U)$ est un espace vectoriel, appelé le k-ième groupe de cohomologie de de Rham. La suite d'homomorphismes (7.56) s'appelle le ***complexe de de Rham*** de U ; en général, une suite d'homomorphismes entre les groupes abéliens comme (7.56) avec $\mathrm{dd} = 0$ s'appellent un ***complexe***.

Le théorème de Poincaré s'énonce maintenant comme suit : $H^k(U) = \{0\}$, $k \geqslant 1$, *si U est une partie ouverte étoilée de $\mathbb{R}^N$ (ou, plus généralement, si U est contractile).* Le fait que $H^0(U) = \mathbb{R}$ veut dire que toute fonction f de classe C^∞ dans U telle que $\mathrm{d}f = 0$ dans U est constante ; c'est évidemment vrai si et seulement si U est un domaine (« ouvert et connexe ») de $\mathbb{R}^N$. Nous étudierons la condition exacte (appelée « connexité simple ») sur U telle que $H^1(U) = \{0\}$ dans les paragraphes suivants. En général, pour chaque $k \geqslant 0$, $\dim H^k(U)$ représente une propriété topologique spécifique de U ; on peut démontrer que si U et V sont deux parties ouvertes de $\mathbb{R}^N$ homéomorphes, alors $\dim H^k(U) = \dim H^k(V)$, $k \geqslant 0$; c'est un cas particulier d'un théorème général important de de Rham. Le théorème est assez facile à démontrer si U et V sont C^∞-difféomorphes mais le cas général est difficile ; il est même un peu surprenant car $H^k(U)$ est défini avec les constructions C^∞ alors qu'il s'avère invariant sous homéomorphismes non nécessairement différentiables.

Les considérations précédentes peuvent être généralisées facilement à une variété différentielle X de classe C^∞ quelconque et les groupes $H^k(X)$ ont les propriétés analogues à celles signalées ci-dessus. Le livre de Bott et Tu [3] donne un exposé moderne de ce sujet.

Pour illustrer en partie la discussion précédente, considérons $U = \mathbb{R}^2 \setminus \{(0,0)\}$; ici $H^0(U) = \mathbb{R}$ car U est connexe ; il n'est pas difficile de démontrer que $H^1(U) = \mathbb{R}$. En effet, on a déjà vu la 1-forme

$$\alpha = \frac{-y}{x^2+y^2}\,\mathrm{d}x + \frac{x}{x^2+y^2}\,\mathrm{d}y$$

plusieurs fois comme une forme fermée ($\mathrm{d}\alpha = 0$) mais non exacte ; l'existence d'une telle forme $\alpha \neq 0$ montre que $\dim H^1(U) \geqslant 1$; il suffit maintenant de voir que si ω est une 1-forme fermée de classe C^∞ dans U, alors $\alpha - \omega$ est exacte ; c'est essentiellement l'exercice 13 (§ 6.2.11). On peut démontrer que $H^2(U) = \{0\}$; pour voir ceci directement, il faut montrer que si $f \in C^\infty(U)$ alors il existe P, Q dans $C^\infty(U)$ telles que

$$f = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \quad \text{dans } U ;$$

en effet, une 2-forme ω dans $\Omega^2(U) = Z^2(U)$ est de la forme

$$\omega = f\,\mathrm{d}x \wedge \mathrm{d}y$$

et il faut établir l'existence de $\pi \in \Omega^1(U)$ telle que $\mathrm{d}\pi = \omega$; si $\pi = P\,\mathrm{d}x + Q\,\mathrm{d}y$ avec $\mathrm{d}\pi = \omega$, il faut avoir P, Q comme indiqué. Nous ne complétons pas cette démarche. En fait, on peut démontrer généralement que si X est une variété différentielle orientée, connexe, non compacte de dimension n alors $H^n(X) = \{0\}$; en particulier, $H^N(U) = \{0\}$ si U est une partie ouverte connexe non compacte de $\mathbb{R}^N$.

Le théorème de Poincaré garantit que toute forme fermée ω de classe C^1 dans une partie ouverte $U \subset \mathbb{R}^N$ est ***localement exacte*** dans le sens que pour chaque point $a \in U$ il existe une boule ouverte $B = B(a;r) \subset U$ telle que $\omega = \mathrm{d}\pi_B$ dans B pour une forme π_B définie dans $B(a;r)$; ceci est dû au fait qu'une boule ouverte (comme tout ensemble convexe) dans $\mathbb{R}^N$ est étoilée ; le problème est de recoller les π_B pour obtenir une forme π définie globalement dans U telle que $\mathrm{d}\pi = \omega$ dans U. On sait que ceci n'est pas toujours possible.

7.10.6 Les 1-formes

Nous considérons maintenant les 1-formes différentielles définies dans une partie ouverte non vide $U \subset \mathbb{R}^N$. En écrivant $x_1, \ldots, x_N$ pour les fonctions coordonnées, une 1-forme ω dans U est de la forme

$$\omega = f_1\,\mathrm{d}x_1 + \cdots + f_N\,\mathrm{d}x_N$$

où $f_1, \ldots, f_N$ sont des fonctions réelles définies sur U. En identifiant $f = (f_1, \ldots, f_N)$ comme un champ vectoriel dans U on peut traduire tous les résultats concernant les intégrales curvilignes donnés dans la section 6.1 comme des énoncés au sujet des 1-formes dans U ; cette possibilité est évoquée explicitement dans (6.3) où on définit $\int_\gamma \omega$, γ étant un chemin ou trajectoire C^1 par morceaux dans U. Nous admettons comme connus tous les résultats de la section 6.1 dans cette notation. Rappelons néanmoins la proposition 6.1.11 dans la terminologie des formes. *Une 1-forme ω continue dans U est exacte (c'est-à-dire il existe une fonction $u \in C^1(U)$ telle que $\mathrm{d}u = \omega$ dans U) si et seulement si*

$$\int_\gamma \omega = 0$$

pour tout lacet γ, C^1 par morceaux, dans U.

On peut même se borner aux lacets polygonaux simples (§ 6.1.12). Si $\mathrm{d}u = \omega$, on a appelé u un potentiel scalaire du champ associé à ω et nous pouvons dire ici simplement un potentiel pour ω ; dans ce cas

$$\int_\gamma \omega = u(q) - u(p)$$

si γ est un chemin quelconque (C^1 par morceaux) reliant p, q dans U. Rappelons que si u est un potentiel pour ω, $u + c$ l'est aussi pour tout $c \in \mathbb{R}$.

Evidemment, si $u \in C^2(U)$ et $\omega = \mathrm{d}u$ alors

$$\mathrm{d}\omega = \mathrm{d}(\mathrm{d}u) = 0$$

et ω est une 1-forme fermée de classe C^1 dans U ; si U est étoilé (ou contractile) on a vu que toute 1-forme fermée ω de classe C^1 dans U dérive d'un potentiel (c'est-à-dire est exacte) : il existe $u \in C^2(U)$ telle que $\mathrm{d}u = \omega$. Dans la section 6.2 on a vu des exemples des 1-formes fermées dans $U \subset \mathbb{R}^2$ qui ne sont pas exactes (c'est-à-dire qui ne dérivent pas d'un potentiel).

La notion d'homotopie des chemins nous permettent d'approfondir ces considérations.

7.10.7 Homotopie des chemins

Soit U une partie ouverte de $\mathbb{R}^N$; soient α, β deux chemins dans U reliant p à q ; sans restriction de généralité, nous pouvons supposer que α, β sont des applications continues de $[0\,,\,1]$ dans U avec

$$\alpha(0) = \beta(0) = p\,, \quad \alpha(1) = \beta(1) = q\,;$$

on aura aucune difficulté à généraliser les considérations présentées ici au cas ou α, $\beta : [a\,,\,b] \to U$. On dira que α, β sont ***homotopes*** (***avec extrémités fixes dans*** U) s'il existe une application continue H (dite ***homotopie***) :

$$H : [0\,,\,1] \times [0\,,\,1] \longrightarrow U$$

telle que

$$H(t,0) = \alpha(t)\,, \quad H(t,1) = \beta(t)\,, \quad t \in [0\,,\,1] \tag{7.57}$$

$$H(0,u) = p\,, \quad H(1,u) = q\,, \quad u \in [0\,,\,1]\,. \tag{7.58}$$

La condition (7.57) correspond à la condition d'homotopie entre les applications α, β comme on l'a vu dans le paragraphe 7.10.2 ; (7.58) exprime le fait que tous les chemins «intermédiaires» $t \mapsto H(t,u)$ ($u \in [0\,,\,1]$ fixe) relient p à q dans U. On dit que l'homotopie H décrit une ***déformation continue*** du chemin α en le chemin β, les déformations restant toujours *dans* U.

Donnons un exemple simple pour illustrer cette définition ; soient

$$\alpha(t) = \big(\cos t\pi\,,\,\sin t\pi\big)\,, \quad \beta(t) = \big(\cos t\pi\,,\,-\sin t\pi\big)\,, \quad 0 \leqslant t \leqslant 1\,.$$

Les chemins α, β relient $(1,0)$ à $(-1,0)$ dans $U = \mathbb{R}^2 \setminus \{(0,0)\}$. Ces deux chemins ne sont pas homotopes (avec extrémités fixes) dans U ; un dessin rend cette affirmation évidente mais sa démonstration rigoureuse la plus simple passe par la proposition 7.10.8 qui suit. En effet, la 1-forme

$$\omega = \frac{-y}{x^2+y^2}\,\mathrm{d}x + \frac{x}{x^2+y^2}\,\mathrm{d}x$$

est de classe C^∞ dans U et elle est fermée ; mais un calcul montre que

$$\int_\alpha \omega = \pi\,, \quad \int_\beta \omega = -\pi$$

ce qui exclut la possibilité d'homotopie entre α et β dans U. Notons, par contre, que α, β, en tant que chemins dans $\mathbb{R}^2$, sont homotopes (avec extrémités fixes) dans $\mathbb{R}^2$; en effet,

$$H(t,u) = (1-u)\alpha(t) + u\beta(t)\,, \quad (t,u) \in [0\,,\,1]^2$$

est une homotopie dans $\mathbb{R}^2$. Observer que

$$H\Big(t, \frac{1}{2}\Big) = (\cos t\pi\,,\,0)\,, \quad 0 \leqslant t \leqslant 1\,,$$

est un chemin de $(1,0)$ à $(-1,0)$ qui passe par le point « interdit » $(0,0)$. Ainsi, étant donnés deux chemins α, β dans U (ayant les mêmes extrémités), le fait qu'ils sont homotopes (avec extrémités fixes) dans U reflète tout autant une propriété géométrique de U que les propriétés de α et β.

Nous donnons maintenant une proposition importante du sujet.

7.10.8 Proposition

Soient U une partie ouverte de $\mathbb{R}^N$, p, q deux points dans U et α, β deux chemins C^1 par morceaux dans U reliant p à q. Si α, β sont homotopes (avec extrémités fixes) dans U alors pour toute 1-forme fermée ω de calsse C^1 dans U on a

$$\int_\alpha \omega = \int_\beta \omega\,.$$

DÉMONSTRATION. Soit $H : [0\,,\,1]^2 \to U$ l'homotopie qui donne la déformation continue de α en β où $\alpha, \beta : [0\,,\,1] \to U$ avec $\alpha(0) = \beta(0) = p$, $\alpha(1) = \beta(1) = q$. Puisque H est continue dans l'ensemble compact $[0\,,\,1]^2$, l'ensemble $K = H\big([0\,,\,1]^2\big)$ est une partie compacte de U et H est uniformément continue. Selon le paragraphe 1.9.1, il existe un $\varepsilon > 0$ tel que

$$\big\{x : d(x,K) < \varepsilon\big\} \subset U\,;$$

en particulier, pour tout $(t,u) \in [0\,,\,1]^2$,

$$B\big(H(t,u);\varepsilon\big) \subset U\,.$$

L'uniforme continuité de H donne l'existence d'un $\delta > 0$ tel que

$$\big\|H(t,u) - H(t',u')\big\| \leqslant \frac{\varepsilon}{2} \tag{7.59}$$

dès que $|t-t'| < \delta$, $|u-u'| < \delta$, $(t,u) \in [0\,,\,1]^2$.

Soient $0 = t_0 < t_1 < \ldots < t_n = 1$ tels que $\max_j(t_j - t_{j-1}) < \delta$; pour $u \in [0, 1]$, soit $\gamma_u : [0, 1] \to U$ le chemin polygonal tel que $\gamma_u(t_j) = H(t_j, u)$, $0 \leqslant j \leqslant n$, γ_u traçant un segment lorsque t va de t_{j-1} à t_j, $1 \leqslant j \leqslant n-1$, c'est-à-dire

$$\gamma_u(t) = \lambda\gamma_u(t_j) + (1-\lambda)\gamma_u(t_{j-1})$$

si $t = \lambda t_j + (1-\lambda)t_{j-1}$, $0 \leqslant \lambda \leqslant 1$; puisque

$$\|\gamma_u(t) - \gamma_u(t_{j-1})\| \leqslant \|\gamma_u(t_j) - \gamma_u(t_{j-1})\| < \frac{\varepsilon}{2}$$

pour $t \in [t_{j-1}, t_j]$, on voit bien que γ_u est un chemin dans U.

Nous démontrons maintenant que

$$\int_\alpha \omega = \int_{\gamma_0} \omega, \quad \int_\beta \omega = \int_{\gamma_1} \omega \tag{7.60}$$

et

$$\int_{\gamma_u} \omega = \int_{\gamma_v} \omega \quad \text{si} \quad |u - v| < \delta,\ u, v \in [0, 1]. \tag{7.61}$$

L'égalité entre $\int_\alpha \omega$ et $\int_\beta \omega$ est une conséquence immédiate de (7.60) et (7.61). Puisque $H(t, 0) = \alpha(t)$, $0 \leqslant t \leqslant 1$, on a

$$p_j = H(t_j, 0) = \alpha(t_j) = \gamma_0(t_j), \quad j = 0, 1, \ldots, n ;$$

posons $B_j = B(p_j; \varepsilon)$. On a $B_j \subset U$ et

$$\alpha([t_{j-1}, t_j]) \subset B_j, \quad \gamma_0([t_{j-1}, t_j]) \subset B_j$$

pour $1 \leqslant j \leqslant n$ à cause de (7.59). Puisque ω est une 1-forme fermée, dans B_j ω dérive d'un potentiel f_j (selon le théorème de Poincaré, voir aussi § 7.10.6) ; notons $\alpha_j = \alpha|_{[t_{j-1}, t_j]}$ et par $\gamma_{0j} = \gamma_0|_{[t_{j-1}, t_j]}$; alors

$$\int_{\alpha_j} \omega = f_j(p_j) - f_j(p_{j-1}) = \int_{\gamma_{0j}} \omega$$

d'où en sommant sur $j = 1, \ldots, n$ on a $\int_\alpha \omega = \int_{\gamma_0} \beta$. La preuve pour l'égalité $\int_\beta \omega = \int_{\gamma_1} \omega$ est parfaitement analogue ; ceci établit (7.60).

Démontrons maintenant que $u \mapsto \int_{\gamma_u} \omega$ est localement constante pour $u \in [0, 1]$ comme précisé dans (7.61). Fixons un $u \in [0, 1]$; posons B_j pour la boule $B(\gamma_u(t_j); \varepsilon)$; soit $f_j : B_j \to \mathbb{R}$ un potentiel pour ω dans B_j. Notons que $B_j \subset U$, $j = 0, \ldots, n$, et que $B_{j+1} \cap B_j$ est non vide et connexe, $j = 0, \ldots, (n-1)$; donc il existe une constante réelle c_j telle que dans $B_{j+1} \cap B_j$ on a

$$f_{j+1} = f_j + c_j$$

car dans le domaine $B_{j+1} \cap B_j$, on a $\mathrm{d}(f_{j+1} - f_j) = 0$. Supposons maintenant que $v \in [0, 1]$ avec $|u - v| < \delta$; alors $\gamma_v(t_{j-1})$ et $\gamma_v(t_j)$ appartiennent à B_j (à cause de (7.59)). En écrivant $\gamma_{vj} = \gamma_v|_{[t_{j-1}, t_j]}$ on a

$$\int_{\gamma_{vj}} \omega = f_j\big(\gamma_v(t_j)\big) - f_j\big(\gamma_v(t_{j-1})\big)$$

d'où, en faisant la somme sur $j = 1, \ldots, n$, on a

$$\int_{\gamma_v} \omega = \sum_{j=1}^{n} \big\{ f_j\big(\gamma_v(t_j)\big) - f_j\big(\gamma_v(t_{j-1})\big) \big\} .$$

Puisque $\gamma_v(t_0) = p$, $\gamma_v(t_n) = q$, le second membre s'écrit comme

$$-f_1(p) + \sum_{j=1}^{n-1} \big\{ f_j\big(\gamma_v(t_j)\big) - f_{j+1}\big(\gamma_v(t_j)\big) \big\} + f_n(q) ;$$

notons que $\gamma_v(t_j) \in B_{j+1} \cap B_j$, $j = 1, \ldots, n-1$; donc on a

$$\int_{\gamma_v} \omega = \big\{ f_n(q) - f_1(p) \big\} - \sum_{j=1}^{n-1} c_j$$

une quantité indépendante de $v \in [0, 1]$ avec $|u - v| < \delta$. Ceci établit (7.61) et achève la démonstration. ♦

Nous déduisons maintenant une conséquence importante de la proposition 7.10.8. Pour cela nous introduisons d'abord une notion essentielle.

7.10.9 Connexité simple

Soit U un domaine de $\mathbb{R}^N$. Si p est un point de U, l'application constante $\beta(t) \equiv p$, $a \leqslant t \leqslant b$, s'appelle un ***lacet ponctuel*** au point p ; comme toujours, il nous suffira, en général, de prendre $a = 0$, $b = 1$.

On dit que U est ***simplement connexe*** s'il existe un point $p \in U$ (appelé ***point base***) tel que tout lacet $\alpha : [0, 1] \to U$, $\alpha(0) = \alpha(1) = p$, est homotope (avec extrémités fixes) au lacet ponctuel $\beta(t) \equiv p$, $0 \leqslant t \leqslant 1$.

Si U est une partie ouverte de $\mathbb{R}^N$ étoilée par rapport à un point $p \in U$, alors U est simplement connexe. En effet, il est clair que U est connexe car si x et y sont deux points de U, le chemin $[x, p] \vee [p, y]$ relie x à y dans U. Soit $\alpha : [0, 1] \to U$ un lacet avec $\alpha(0) = \alpha(1) = p$. Posons

$$H(t, u) := up + (1 - u)\alpha(t) , \quad (t, u) \in [0, 1]^2$$

alors

$$H(t, 0) = \alpha(t) , \quad H(t, 1) = p , \quad 0 \leqslant t \leqslant 1$$

et

$$H(0,u) = H(1,u) = p\,, \quad 0 \leqslant u \leqslant 1\,.$$

Donc H est l'homotopie désirée entre α et $\beta \equiv p$, établissant la connexité simple de U.

7.10.10 Proposition

Soit U un domaine simplement connexe dans $\mathbb{R}^N$; alors toute 1-forme fermée ω de classe C^1 dans U est exacte.

DÉMONSTRATION. Il suffit de montrer que

$$\int_\alpha \omega = 0$$

pour tout lacet α, C^1 par morceaux, dans U. Supposons d'abord que $\alpha : [0\,,1] \to U$ avec $\alpha(0) = \alpha(1) = p$ où p est le point base dans la définition de la connexité simple de U ; puisque α est homotope (avec extrémités fixes) au lacet ponctuel $\beta(t) \equiv p$, $0 \leqslant t \leqslant 1$, il résulte de la proposition 7.10.8 que

$$\int_\alpha \omega = \int_\beta \omega = 0\,.$$

Si $\alpha : [0\,,1] \to U$ est un lacet, C^1 par morceaux, avec $\alpha(0) = \alpha(1) = q \neq p$, nous prenons un chemin β, C^1 par morceaux, reliant p à q dans U (β existe car U est ouvert et connexe, donc, connexe par arcs ; prop. 1.8.2) et posons

$$\gamma(t) = \begin{cases} \beta(t) & \text{si } 0 \leqslant t \leqslant 1 \\ \alpha(t-1) & \text{si } 1 \leqslant t \leqslant 2 \\ \beta(3-t) & \text{si } 2 \leqslant t \leqslant 3\,; \end{cases}$$

autrement dit, $\gamma = \beta \vee \alpha \vee \widetilde{\beta}$. Il est clair que $\gamma : [0\,,3] \to U$ est un lacet, C^1 par morceaux et tel que $\gamma(0) = \gamma(3) = p$; selon le raisonnement déjà présenté,

$$\int_\gamma \omega = 0\,.$$

Or,

$$\int_\gamma \omega = \int_\beta \omega + \int_\alpha \omega - \int_\beta \omega = \int_\alpha \omega$$

d'où $\int_\alpha \omega = 0$. ♦

7.10.11 Remarques

Il n'est pas très facile, en général, de démontrer explicitement si un domaine donné de $\mathbb{R}^N$ est simplement connexe ou non ; souvent, il est géométriquement évident qu'un domaine est simplement connexe mais la démonstration

directe peut être fastidieuse et il est presque toujours difficile de démontrer directement qu'un domaine n'est pas simplement connexe. Pour ces raisons, nous donnons quelques indications sans preuves qui peuvent être utiles dans ce contexte.

On a déjà vu que les domaines étoilés de $\mathbb{R}^N$ sont simplement connexes. Il n'est pas difficile d'établir que si U et V sont deux domaines homéomorphes de $\mathbb{R}^N$ alors V est simplement connexe si U est simplement connexe ; ainsi, tout domaine homéomorphe à un domaine étoilé est simplement connexe.

Si un domaine est simplement connexe par rapport à un point base, il est simplement connexe par rapport à tout autre point base. Un domaine U est simplement connexe si et seulement si deux chemins quelconques dans U, ayant même origine et même fin, sont homotopes avec extrémités fixes dans U.

Un domaine contractile est simplement connexe. L'ensemble $U = \mathbb{R}^N \setminus \{0\}$ est simplement connexe si $N \geqslant 3$; si $N = 1$, U est ouvert mais non connexe ; si $N = 2$, U est un domaine non simplement connexe. La preuve la plus simple de cette dernière affirmation utilise la proposition 7.10.10 ; puisque U possède une 1-forme fermée ω (de classe C^∞) non exacte, U ne peut pas être simplement connexe ; il suffit de prendre

$$\omega = \frac{-y}{x^2+y^2}\,\mathrm{d}x + \frac{x}{x^2+y^2}\,\mathrm{d}y\,.$$

Avec la notation cohomologique introduite dans le paragraphe 7.10.5 on peut dire que, si U est un domaine simplement connexe dans $\mathbb{R}^N$, alors $H^1(U) = \{0\}$; en fait, la réciproque est vraie aussi : si U est un domaine dans $\mathbb{R}^N$ tel que $H^1(U) = \{0\}$ alors U est simplement connexe.

Notons que pour la connexité simple une définition d'***homotopie de lacets*** est plus commode. Soit U un domaine de $\mathbb{R}^n$; deux lacets α, β sont ***homotopes comme lacets*** dans U (avec $\alpha, \beta : [0\,,\,1] \to U$ continus, $\alpha(0) = \alpha(1)$, $\beta(0) = \beta(1)$) s'il existe une homotopie $H : [0\,,\,1]^2 \to U$ de α à β (donc H est continue et $H(t,0) = \alpha(t)$, $H(t,1) = \beta(t)$, $t \in [0\,,\,1]$) telle que $H(0,u) = H(1,u)$, $0 \leqslant u \leqslant 1$; ceci veut dire que pour tout u, le chemin $t \mapsto H(t,u)$ est un lacet. On peut démontrer qu'un domaine U est simplement connexe (selon notre définition utilisant un point base fixe) si et seulement si tout lacet dans U est homotope comme lacet dans U à un lacet ponctuel. Pour les lacets, c'est la notion d'homotopie comme lacet (appelée aussi homotopie libre) qui est la plus commode.

Pour avoir une théorie plus satisfaisante, on doit étudier le ***groupe fondamental*** $\pi_1(U)$ associé à un domaine U (plus généralement, à un espace topologique U). Sans définir $\pi_1(U)$ disons simplement que $\pi_1(U)$ est un groupe (en général, non commutatif) associé à U tel que $\pi_1(U) = \{e\}$ (groupe contenant seulement l'élément identité) si et seulement si U est simplement connexe. L'étude de la connexité simple de U revient donc à la considération du groupe $\pi_1(U)$ pour lequel il y a des méthodes de calcul précises données dans les livres de topologie algébrique.

7.10.12 Exercices

1 Soient α, β deux k-formes ($k \geqslant 1$) de classe C^1 dans un domaine étoilé U de $\mathbb{R}^N$; si $\mathrm{d}\alpha = \mathrm{d}\beta$ dans U, démontrer que $\beta = \alpha + \mathrm{d}\theta$ dans U où θ est une $(k-1)$-forme de classe C^1 dans U.

2 Soient α, β deux formes de classe C^1 dans une partie ouverte U de $\mathbb{R}^N$; si α est fermée et β est exacte démontrer que $\alpha \wedge \beta$ est exacte.

3 Soit $\varphi : U \to V$ un C^2-difféomorphisme entre les deux parties ouvertes U, V de $\mathbb{R}^N$. Vérifier que via φ^* on peut établir un isomorphisme entre les groupes $H^k(U)$ et $H^k(V)$, $k = 0, \ldots, N$.

Conclusion

Pour aller plus loin dans l'étude de l'analyse dans les variétés, il faut acquérir la connaissance de base concernant la topologie générale et l'analyse fonctionnelle. Dieudonné [13, t. 1] et Schwartz [38, t. 1] sont des bonnes références à ce sujet. Les fondements de l'analyse moderne sont présentés d'une manière succincte mais accessible dans les deux ouvrages de Rudin [36] et [37] ; pour pouvoir lire avec profit une référence comme [13] il est nécessaire d'avoir la préparation solide fournie par [36].

Les variétés considérées dans notre texte sont les sous-ensembles d'un $\mathbb{R}^N$; il est utile de se familiariser avec la notion abstraite d'une variété en tant qu'un certain type d'espace topologique ; Dieudonné [13, t. 3] donne une introduction détaillée à ce sujet. Pour comprendre l'importance de cet aspect des choses, considérons la notion d'un espace vectoriel réel ; on peut la définir abstraitement, comme c'est habituel aujourd'hui, en donnant les axiomes d'un espace vectoriel réel ; mais on peut aussi dire que c'est un espace de fonctions réelles définies dans un ensemble I. Il est vrai que tout espace vectoriel réel est isomorphe à un espace des fonctions réelles définies dans un ensemble approprié I, mais tout le monde est d'accord aujourd'hui de dire que la définition axiomatique d'un espace vectoriel réel est plus satisfaisante. Dans la théorie des variétés de dimension finie, on démontre, après coup, que chacune de ces variétés est homéomorphe (ou difféomorphe) à une sous-variété d'un espace $\mathbb{R}^N$ mais l'intérêt d'une définition axiomatique d'une variété ne s'en trouve pas diminué pour autant.

Curieusement, l'axiomatisation de la notion de variété est relativement récente, comme d'ailleurs l'est aussi l'axiomatisation de la notion d'espace vectoriel. Ces axiomatisations, introduites dans les années 1930, deviennent généralement connues seulement après 1945, même si leurs structures étaient clairement comprises depuis le milieu du siècle précédent au moins. Citons un passage de Poincaré dans « Les méthodes nouvelles de la mécanique céleste » (tome 3, 1899, p. 6) pour illustrer l'esprit d'une certaine époque : « Le mot *variété* est maintenant assez usité pour que je n'aie pas cru nécessaire d'en rappeler la définition. On appelle ainsi tout ensemble continu de points (ou de système de valeurs) : c'est ainsi que dans l'espace à trois dimensions, une surface quelconque est une variété à deux dimensions et une ligne quelconque, une variété à une dimension. » Cette description imprécise n'empêche pas Poincaré de développer une analyse puissante dans les variétés, contenant une version du théorème de Stokes formulée en termes des formes différentielles et bien d'autres choses ; on peut même nommer Poincaré le fondateur de l'analyse dans les variétés.

Pour les questions historiques d'ordre général concernant les mathématiques modernes, les ouvrages de Bourbaki [5] et Dieudonné [14] sont utiles ; le livre de Crowe [10] est une étude détaillée de l'histoire de l'analyse vectorielle élémentaire. Une histoire détaillée de l'analyse dans les variétés n'est pas encore faite ; elle doit remonter aux travaux de Gauss, peut être même à ceux d'Euler et de Lagrange et certainement à une contribution remarquable de Riemann « *Über die Hypothesen, welche der Geometrie zu Grunde liegen* », lu en 1854 à l'occasion de ses épreuves d'admission (*Habilitationsvortrag*) à la Faculté philosophique de Göttingen et publié en 1867 après sa mort en 1866. La monographie de Dieudonné « *A history of algebraic and differential topology* 1900-1960 » (Birkhäuser, 1989) fait une mise au point très lisible dans son chapitre III de la première partie (pp. 60-66) ; une autre référence est celle de E. Scholz « *Geschichte der Mannigfaltigkeitsbegriffs von Riemann bis Poincaré* » (Birkhäuser, 1980). Une référence générale, surtout pour l'information concernant la vie et l'œuvre des grands scientifiques, est « *Dictionary of scientific biography* » (D.S.B.), éditeur C. C. Gillispie, Scribner's, N. Y. 1981 (en 8 volumes).

L'analyse dans les variétés est un sujet central des mathématiques ; on doit mentionner encore l'analyse dans les variétés complexes qui sont des sous-ensembles de $\mathbb{C}^N$ comme nos variétés dans ce texte sont des sous-ensembles de $\mathbb{R}^N$. Outre l'ouvrage de Dieudonné [13, t. 3], nous nous référons au livre de Narasimhan [29] comme une introduction pour les étudiants avancés à l'analyse dans les variétés réelles ou complexes. L'ouvrage classique sur l'analyse dans les variétés réelles est celui de de Rham [34] ; ce livre doit être considéré aujourd'hui comme celui qui définit les frontières essentielles de l'analyse dans les variétés réelles. Pour le théorème de Stokes un exposé classique est celui de Whitney [43], un des fondateurs modernes de la « topologie différentielle », réelle ou complexe. Dans le livre de de Rham, on trouvera les références détaillées aux travaux de l'autre fondateur moderne de l'analyse dans les variétés, E. Cartan. Une formulation définitive de la théorie de l'intégration dans les variétés et ses applications au calcul des variations se trouve dans le livre de Federer [18] auquel nous nous sommes référé plusieurs fois ; malheureusement, sa lecture est une tâche ardue ; il serait souhaitable d'avoir un jour un accès plus facile aux résultats remarquables contenus dans le livre de Federer.

Réponses aux exercices

CHAPITRE 1

1.1.1

[5] Soit $f_n(x) = n - n^2x$ si $0 \leqslant x \leqslant 1/n$, $f_n(x) = 0$ si $1/n < x \leqslant 1$, $n = 2, 3, \ldots$; on a $\|f_n\|_\infty = n$, $\|f_n\|_1 = 1$.

1.4.7

[3] $\overline{A} = \{(x,y) : 0 < y \leqslant 1/x\} \cup \{(x,0) : x \geqslant 0\} \cup \{(0,y) : y \geqslant 0\}$.

1.7.4

[3] $A =]0, 1/2] \subset X =]0, 1[$; A est borné, X-fermé, non compact.

[4] $\exists R > 0$ t.q. $f(x) < M/2$ si $\|x\| > R$; donc $\exists \xi$ t.q. $\|\xi\| \leqslant R$ avec $f(\xi) = M$.

[6] $\forall \varepsilon > 0$, $\exists \delta > 0$ t.q. $f(x,y) - \varepsilon < f(x',y) < f(x,y) + \varepsilon$ si $y \in J_2$ et $|x - x'| < \varepsilon$ (de la continuité uniforme de f) ; donc $\varphi(x) - \varepsilon \leqslant \varphi(x') \leqslant \varphi(x) + \varepsilon$ si $|x - x'| < \varepsilon$.

1.8.4

[6] Soit $S^{N-1} = \{x \in \mathbb{R}^N : \|x\| = 1\}$, $N \geqslant 2$; l'ensemble $D_n = \{x \in \mathbb{R}^n : \|x\| < 1\}$ est convexe donc connexe ; l'ensemble $S_+^{N-1} = \{x \in S^{N-1} : x_N > 0\}$, $x = (x_1, \ldots, x_N)$, est connexe car l'image de D_{N-1} par la fonction continue

$$(x_1, \ldots, x_{N-1}) \mapsto \left(x_1, \ldots, x_{N-1}, \sqrt{1 - x_1^2 - \cdots - x_{N-1}^2}\right) ;$$

de même pour S_-^{N-1} ; donc leurs adhérences sont connexes ; S^{N-1} est égal à $A \cup B$, $A = \overline{S}_+^{N-1}$, $B = \overline{S}_-^{N-1}$, $A \cap B \neq \emptyset$ donc S^{N-1} est connexe.

1.9.6

[4] $1/\sqrt{2}$.

[6] Noter que $\sup_x g_n(x) = 1/4$.

7 Si $x \in \overline{V}$ il existe $\{x_n\}$ dans V t.q. $x_n \to x$; la continuité uniforme de f donne que $\{f(x_n)\}$ est une suite de Cauchy ; donc $\widetilde{f}(x) = \lim_n f(x_n)$ existe, etc.

1.10

2 $A =]0, 1]$, $B = \{0\}$; $A \cap B = \emptyset$ mais $\overline{A} \cap \overline{B} = \{0\}$.

3 $\mathbb{R}^N$ étant connexe, soit A, soit A^c n'est pas fermé ; si A n'est pas fermé, $\exists\, a \in \overline{A} \setminus A$ et $a \in \operatorname{Fr} A$, etc.

4 $\operatorname{Fr} F = F$.

5 $\operatorname{Fr} A = \emptyset \Leftrightarrow A = \emptyset$ ou $\mathbb{R}^N$; utiliser la connexité de $\mathbb{R}^N$: si $\operatorname{Fr} A = \emptyset$, $\mathbb{R}^N = \mathring{A} \cup \operatorname{ext} A$, etc. De même, $\operatorname{Fr} A = \{a\} \Leftrightarrow A = \{a\}$ ou $\{a\}^c$ si $N \geqslant 2$; pour $\mathbb{R}^1$, $\operatorname{Fr} A = \{a\} \Leftrightarrow A =]-\infty, a[$ ou $]-\infty, a]$ ou $]a, \infty[$ ou $[a, \infty[$. Soit $S = [a, b]$; alors $\operatorname{Fr} A = S \Leftrightarrow A \subset S$ et $\overline{A} = S$ ou $A^c \subset S$ et $\overline{A^c} = S$ si $N \geqslant 2$; pour $\mathbb{R}^1$, $\operatorname{Fr} A = S$ est impossible. Noter que $\mathbb{R}^N \setminus S$ est connexe si $N \geqslant 2$ et $\mathbb{R}^N = \mathring{A} \cup S \cup \operatorname{ext} A$.

6 f est continue.

9 Démontrons que f est continue en $(0,0)$ en supposant que $f(0,0) = 0$; la continuité de $x \mapsto f(x,0)$ et de $y \mapsto f(\pm h_1, y)$ donne successivement $h_1 > 0$ t.q. $|f(x,0)| < \varepsilon/2$ pour $|x| \leqslant h_1$ et $h_2 > 0$ t.q. $|f(\pm h_1, y) - f(\pm h_1, 0)| \leqslant \varepsilon/2$ si $|y| \leqslant h_2$. Si $|x| < h_1$, $|y| < h_2$ alors

$$f(x,y) \leqslant f(h_1,y) \leqslant f(h_1,h_2) < f(h_1,0) + \frac{\varepsilon}{2} < \varepsilon$$

et, de même, $f(x,y) > -\varepsilon$.

11 Il suffit de voir que $u(x) \neq 0$ si $x \neq 0$; si $m = \inf_{\|x\|=1} p\big(u(x)\big)$ alors u^{-1} existe $\Leftrightarrow m > 0$; aussi $p\big(u(x)\big) \geqslant m\|x\|$, $x \in \mathbb{R}^N$ ce qui donne le résultat (Ex. 5, § 1.7.4).

12 $f^{-1}(t) = \big(2t/(t^2+1), (t^2-1)/(t^2+1)\big)$, $t \in \mathbb{R}$.

18 $f(x) = \inf_{a \in A} \|x - a\|$; $\exists\, a_n \in A$ t.q. $\|x - a_n\| \to f(x)$ $(n \to \infty)$; $\{a_n\}$ est bornée car $\|a_n\| \leqslant \|a_n - x\| + \|x\|$; donc, $\exists \{a_{n_j}\}$, $a_{n_j} \to a \in A$ avec $\|x - a\| = f(x)$. Pour obtenir l'unicité de a lorsque A est convexe utiliser l'inégalité

$$\left\| x - \frac{a+b}{2} \right\|^2 < \frac{1}{2}\left(\|x-a\|^2 + \|x-b\|^2 \right), \quad x \in \mathbb{R}^N,\ a \neq b \in \mathbb{R}^N,$$

conséquence de l'égalité dite du parallélogramme

$$2\left(\|\xi\|^2 + \|\eta\|^2 \right) = \|\xi - \eta\|^2 + \|\xi + \eta\|^2, \quad \xi, \eta \in \mathbb{R}^N.$$

19 Si $\{\xi_1, \ldots, \xi_M\}$, $\{\eta_1, \ldots, \eta_N\}$ sont des bases canoniques de $\mathbb{R}^M$ et $\mathbb{R}^N$ respectivement, posons $B(\xi_i, \eta_j) = (b_{ij}^1, \ldots, b_{ij}^P)$; alors, si $x = (x_1, \ldots, x_N)$, $y = (y_1, \ldots, y_M)$, on vérifie que

$$B(x,y) = \left(\sum_{i,j} b_{ij}^k x_i y_j \right)_{1 \leqslant k \leqslant P}$$

d'où

$$|B(x,y)| \leqslant \alpha |x|\,|y|\,, \quad \alpha = \max_k \sum_{i,j} |b_{ij}^k|\,;$$

on conclut avec l'exercice 5 (§ 1.7.4).

CHAPITRE 2

2.1.6

1 Les gradients sont

(i) $\frac{1}{\sqrt{1-x^2y^2}}(x,y)$;

(ii) $-\sin(x+y+z)(1,1,1)$;

(iii) $\big(2x\cos(x^2+y^2),\ 2y\cos(x^2+y^2),\ -\sin z\big)$;

(iv) $(yx^{y-1},\ x^y \ln x)$;

(v) $\dfrac{1}{x+y\cos z}(1,\ \cos z,\ -y\sin z)$;

(vi) $x/\|x\|$;

(vii) $-2\,\mathrm{e}^{-\|x\|^2}x$;

(viii) c.

2 (i) $r(-\sin t, \cos t)$; (ii) $(-r\alpha \sin \alpha t,\ r\alpha \cos \alpha t,\ \beta)$.

3 (i) 0 ; (ii) 0.

4 Pour (i), (ii), (iii) et (iv) respectivement :

$$\begin{bmatrix} y & x \\ 1 & 1 \\ 1 & -1 \end{bmatrix}\ ; \qquad \begin{bmatrix} yz & xz & xy \\ 1 & 1 & 1 \end{bmatrix}\ ; \qquad A\ ;$$

et

$$\begin{bmatrix} 2x_1 & 2x_2 & \ldots & 2x_N \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & 1 \end{bmatrix} \text{ matrice } (N+1, N)\,.$$

7 $D_1 f(0,0) = 0$, $D_2 f(0,0) = 0$; $f(x,y) = \sqrt{x^2+y^2}\rho(x,y)$ où $\rho(x,y) = \varphi(r)/r \to 0$ lorsque $r \to 0$; donc f est dérivable en $(0,0)$ avec $f'(0,0) = (0,0)$.

11 Noter que f', g' et φ' sont des applications linéaires de $\mathbb{R}^N$ dans $\mathbb{R}^3$; on a

$$\begin{aligned}\varphi(a+h) &= \{f(a) + f'(a)h + \|h\|\rho_1(h)\} \times \{g(a) + g'(a)h + \|h\|\rho_2(h)\} \\ &= \varphi(a) + \{f'(a)h \times g(a)\} + \{f(a) \times g'(a)h\} + R\end{aligned}$$

avec $(R/\|h\|) \to 0$ lorsque $\|h\| \to 0$.

2.2.4

1 $(g \circ f)'(0,0,0) = g'(-1,1)f'(0,0,0)$

$$= \begin{bmatrix} 2 & -3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 6 & 2 \\ 1 & 6 & 2 \end{bmatrix}.$$

2 $g'(0)$, $f'(0,1,0,1)$ $f(0,1,0,1)$, $g'(0)$ $f'(0,1,0,1)$ et $f'(0,1,0,1)g'(0)$ respectivement :

$$\begin{bmatrix} 0 \\ -1 \\ 0 \\ 3 \end{bmatrix}, \qquad [-5,0,4,0], \qquad 0, \qquad \begin{bmatrix} 0 & 0 & 0 & 0 \\ 5 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 \\ -15 & 0 & 12 & 0 \end{bmatrix}, \qquad 0.$$

3 $f'(0,0,1)$, c, $(f \circ g)'(0,0)$ respectivement :

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & c & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \operatorname{arc\,tg} \frac{\pi}{4}, \qquad f'(0,0,1)g'(0,0) = \begin{bmatrix} 0 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

4 On trouve

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

5 Il vient

$$(g \circ f)'(0,0) = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}, \quad (f \circ g)'(0,0,0) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

2.4.5

[1]

(i) aucun point critique ;

(ii) $(0,0)$, minimum absolu ;

(iii) $(0,0)$, n'est pas un extremum relatif ;

(iv) $(-1,1)$, n'est pas un extremum relatif ;

(v) $(0,0)$; $A = \{(x,y) : x^2 + y^2 = 1\}$; $(0,0)$ est un minimum absolu ; chaque point de A est un point de maximum absolu ;

(vi) $(2,-2)$, miminum absolu ; noter que $f = (x+y)^2 + (y+2)^2 - 2$.

2.7

[2] (i) $-(3+\sqrt{3})/(2\sqrt{13})$; (ii) 0 ; f est décroissante au point $(3,2)$ dans la direction de (i) ; f possède un maximum local au point $(3,2)$ dans la direction de (ii) ; pour ce dernier étudier $f(3-2t,\, 2+3t)$ comme fonction de t.

[5] $1/\sqrt{2}$.

[6] $(\cos\alpha + \cos\beta + \cos\gamma)/3$.

[14] $u(x,y) = x^2/2 + y^2 \ln x + \sin y - 1/2$ $(x > 0,\ y \in \mathbb{R})$; u est uniquement déterminé dans $\{(x,y) : x > 0,\ y \in \mathbb{R}\}$ mais dans $\{(x,y) : x \neq 0,\ y \in \mathbb{R}\}$ u n'est pas uniquement déterminé ; pour $x < 0$, $u(x,y) = x^2/2 + y^2 \ln x + \alpha(y)$, avec $\alpha : \mathbb{R} \to \mathbb{R}$ arbitraire, donne une solution.

[15] Un exemple est :

$$f(x,y) = \begin{cases} x^2 & \text{si } x < 0 \text{ ou si } x > 0,\ y > 0 \\ 0 & \text{si } x > 0,\ y < 0 \text{ ou si } x = 0,\ y \neq 0. \end{cases}$$

CHAPITRE 3

3.2.6

[2]

(a) $D_1^k D_2^{n-k}(x\,e^y) = x\,e^y$ si $k = 0$, ou e^y si $k = 1$, ou 0 si $k \geqslant 2$.

(b) $D_1^{n-k} D_2^k(y^2/x) = (-1)^n n!\, x^{-(n+1)} y^2$ si $k = 0$, ou $2(-1)^{n-1}(n-1)!\, x^{-n} y$ si $k = 1$, ou 0 si $k \geqslant 2$.

(c) $D_1^k D_2^{n-k}\big((xy)^{-1}\big) = (-1)^n k!\,(n-k)!\, x^{-(k+1)} y^{-(n-k+1)}$.

3.3.5

1

(i) Polynôme de Taylor d'ordre 4 pour $\cos x \cos y$ à $(0,0)$:

$$1 - \left(\frac{x^2}{2!} + \frac{y^2}{2!}\right) + \left(\frac{x^2y^2}{(2!)^2} + \frac{x^4}{4!} + \frac{y^4}{4!}\right).$$

(ii) Polynôme de Taylor d'ordre 3 pour $e^x \sin y$ à $(0,0)$:

$$y + xy + \left(\frac{x^2y}{2!} - \frac{y^3}{3!}\right).$$

4 Polynôme de Taylor d'ordre n de f au point (x_0, y_0) :

$$\sum_{k=0}^{n} \frac{a^{(k)}(t_0)}{k!}(x - x_0 + y - y_0)^k, \quad t_0 = x_0 + y_0 ;$$

le coefficient de $(x - x_0)^i (y - y_0)^{k-i} = a^{(k)}(t_0)/(i!(k-i)!)$.

5 $f = -3 - 2\{(x-1) + (y-1) + (z-1)\} + \{(x-1)^2 + (y-1)^2 + (z-1)^2 - 2(x-1)(y-1) - 2(x-1)(z-1) - 2(y-1)(z-1)\}$.

6

(i) $-(x-1)^2 2y^2 + (2/3)(x-1)^3 - 4(x-1)y^2$; écrire $x^2 + 2y^2 = t$ comme

$$(x-1)^2 + 2(x-1) + 2y^2 + 1 \quad \text{et} \quad \ln t = (t-1) - \frac{(t-1)^2}{2} + \frac{(t-1)^3}{3} + \cdots ;$$

(ii) Ecrire $xy = t$ comme $(x-1)(y-1) + (x-1) + (y-1) + 1$ et

$$e^t = e\left[1 + (t-1) + \frac{1}{2}(t-1)^2 + \frac{1}{6}(t-1)^3 + \cdots\right] ;$$

alors on trouve le polynôme de Taylor :

$$e^t = e\left[1 + (x-1) + (y-1) + \frac{(x-1)^2}{2} + 2(x-1)(y-1) + \frac{(y-1)^2}{2} + \frac{(x-1)^3}{6} + \frac{3}{2}(x-1)^2(y-1) + \frac{3}{2}(x-1)(y-1)^2 + \frac{1}{6}(y-1)^3\right].$$

(iii) $x + yz - \frac{1}{2}xy^2$.

3.4.8

1

(i) Strictement positif car la forme s'écrit $(x_1 - x_2 + 2x_3)^2 + (x_2 - x_3)^2 + x_3^2$ et la matrice est

$$\begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -3 \\ 2 & -3 & 3 \end{bmatrix} ;$$

(ii) indéfini car la forme s'écrit $(2x_1 + x_2 - 2x_3)^2 + x_2^2 - 3x_3^2$ et la matrice est

$$\begin{bmatrix} 4 & 2 & -2 \\ 2 & 2 & -2 \\ -2 & -2 & 1 \end{bmatrix} ;$$

(iii) positif car la forme s'écrit $(2x_1 + 3x_2 + x_3)^2 + x_3^2$; la matrice vaut

$$\begin{bmatrix} 4 & -6 & 2 \\ -6 & 9 & -3 \\ 2 & -3 & 2 \end{bmatrix} ;$$

(iv) indéfini car la forme vaut 1 pour $(1,0,1)$ et -1 pour $(1,0,-1)$;

$$\begin{bmatrix} 0 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 0 \end{bmatrix} ;$$

(v) indéfini car la forme vaut 1 pour $(1,0,0,0)$ et -1 pour $(0,0,0,1)$ et la matrice vaut

$$\begin{bmatrix} 4 & 0 & 0 & -6 \\ 0 & -1 & 3/2 & 0 \\ 0 & 3/2 & 0 & 0 \\ -6 & 0 & 0 & -1 \end{bmatrix} .$$

2 $A \neq 0$ et $\operatorname{tr} A = 0$ entraînent que A possède des valeurs propres de signes opposés d'où le résultat.

4

(i) Si $a = 0$, $(0,0)$ est le seul point critique qui est un minimum absolu. Si $a \neq 0$, les points critiques sont (0,0), $(\pm 2a^{-2}, -a^{-1})$; $(0,0)$ est un point de minimum local strict, les deux autres sont des points de selle ; pour $(0,0)$ il faut faire une analyse directe de $f = x^2(1 + ay) + y^4$ alors que pour les deux autres la matrice hessienne donne la conclusion.

(ii) Points critiques : $(0,0)$, $(1/3)(2^{-1/3}, 2^{-2/3})$; $(0,0)$ est un point de selle ; l'autre est un point de minimum local strict.

(iii) Points critiques : $(0,0)$, $(1,1)$; $(0,0)$ est un point de selle ; $(1,1)$ est un point de minimum local strict.

(iv) Points critiques : $(0,0)$, $(1,1)$, $(-1,-1)$; $(0,0)$ est un point de selle ; les deux autres sont des points de minimum local strict.

(v) $n = 1$: le point critique est $(1,1)$ qui est un point de selle ; $n = 2$: les points critiques sont $\{(x,y) : y = x\}$, points de minimum absolu ; $n \geqslant 3$ impair : comme dans (iii) ; $n \geqslant 4$ pair : comme dans (iv).

(vi) Se référer à l'exercice 4 (§ 2.4.5).

(vii) Points critiques : $(0,0)$, $(1/3,1/3)$; $(0,0)$ est un point de selle ; $(1/3,1/3)$ est un point de maximum relatif strict.

(viii) Points critiques : $(1/\sqrt{2},1/\sqrt{2})$, $(-1/\sqrt{2},-1/\sqrt{2})$; $(1/\sqrt{2},1/\sqrt{2})$ est un point de maximum absolu ; $(-1/\sqrt{2},-1/\sqrt{2})$ est un point de minimum absolu ; une vérification simple de cette dernière affirmation se fait en écrivant f en coordonnées polaires

$$f(r\cos\theta\,,\,r\sin\theta) = \frac{r}{r^2+1}\,(\cos\theta + \sin\theta)\,.$$

CHAPITRE 4

4.1.4

[2] $-2(x-1)-y$.

[3] $1/2$.

[4] $f'(-1) = 2$, $g'(-1) = -3$.

[5] $-(1,1,1)$.

[7] $\nabla f(1,1,1) = (-2,-1,-1)$, $\nabla g(1,1,1) = (0,-1,0)$.

4.2.10

[7] $\varphi^{-1}(u,v) = \big(f(u,v),\, v/\big(2f(u,v)\big)\big)$ où

$$f(u,v) = \left\{\frac{1}{2}\left(\left(u^2+v^2\right)^{1/2}+u\right)\right\}^{1/2}$$

[9] $\varphi^{-1}(u,v) = \big(f(u,v),\, f(u,v)+v\big)$ où $f(u,v) = \big(3(u+v)\big)^{1/3}$.

4.3.3

[3] On trouve

$$\nabla f(1,-1) = \frac{-(2n+1)}{\alpha}(1,1)\,.$$

[4] $y = -x + 1$.

[5] Une condition suffisante est $|f'(1)| \neq |g'(1)|$; on a alors $\varphi'(1) = -1$, $\psi'(1) = 1$ et l'angle d'intersection est $\pi/2$.

[6] On a

$$F'(x,y) = \frac{1}{6fg(2g-f)} \begin{bmatrix} 6g^2/x & -2g/y \\ -3f^2/x & 2f/y \end{bmatrix}$$

où f, g sont calculés au point (x,y) dans un voisinage ouvert de (e^2, e^3) ; noter que $2g - f = 1$ au point (e^2, e^3).

[7]

(a) Une condition suffisante est que

$$\det \begin{bmatrix} f'(x_0) & f'(y_0) & f'(z_0) \\ g'(x_0) & g'(y_0) & g'(z_0) \\ h'(x_0) & h'(y_0) & h'(z_0) \end{bmatrix} \neq 0\,.$$

(b) Une condition suffisante ici est que x_0, y_0 et z_0 soient trois nombres distincts.

(c) Pour n équations la situation est parfaitement analogue.

4.4.10

[8] $\varphi(u,v) = u^2 + v^2 - 1$; $f(\Omega)$ est le cercle unité.

[9] $\varphi(u,v,w) = u^2 + v^2 - 2w$; $f(\Omega)$ est un paraboloïde de révolution.

4.5.3

[1]

(a) $(1/2\,,\,1/2)$, point de maximum global.

(b) $2^{-1/\alpha}(1,1)$, point de maximum global.

(c) $(ab + ac + bc)^{-1}(bc, ac, ab)$, point de minimum global.

(d) $(2,2,1)$, $(2,1,2)$, $(1,2,2)$ sont les points de minimum global ;

$$\left(\frac{4}{3}, \frac{4}{3}, \frac{7}{3}\right), \quad \left(\frac{4}{3}, \frac{7}{3}, \frac{4}{3}\right), \quad \left(\frac{7}{3}, \frac{4}{3}, \frac{4}{3}\right)$$

sont les points de maximum global ;

(e) Le seul point stationnaire $(0,0,0)$ n'est pas un extremum local.

2 Si $(a,b) \neq (0,0)$, l'infimum est atteint au point $\left(1/\sqrt{a^2+b^2}\right)(a,b)$.

3 $\alpha = a/\|a\|$, $f(\alpha) = \left(\|a\| - 1\right)^2$.

4 $M = \sqrt{\alpha^2+\beta^2}$, $\mu = -\sqrt{\alpha^2+\beta^2}$.

CHAPITRE 5

5.4.5

3 Soit $E = \left\{(x,y) \in A \times A : f(x)g(y) - f(y)g(x) \neq 0\right\}$; si $\lambda_N \otimes \lambda_N(E) = 0$, on aura $\lambda_N(E^y) = 0$ p.p. $y \in A$; si g n'est pas 0 p.p., on aura $y_0 \in A$ t.q. $g(y_0) \neq 0$ et $\lambda_N(E^{y_0}) = 0$; puisque $E^{y_0} = \left\{x \in A : f(x) \neq \lambda g(x)\right\}$ avec $\lambda = f(y_0)/g(y_0)$ on conclut que $f(x) = \lambda g(x)$ p.p. $x \in A$.

4 Soit

$$E = \left\{(x,y) \in A \times A : \left[f(x) - f(y)\right]\left[g(x) - g(y)\right] \neq 0\right\} ;$$

si $\lambda_N \otimes \lambda_N(E) = 0$, on aura $\lambda_N(E^y) = 0$ p.p. $y \in A$; donc $\exists\, y_0 \in A$ tel que

$$\left[f(x) - f(y_0)\right]\left[g(x) - g(y_0)\right] = 0 \quad \text{p.p. } x \in A .$$

Posons $B = \left\{x \in A : g(x) = g(y_0)\right\}$, $B' = \left\{x \in A : g(x) \neq g(y_0)\right\}$; si $\lambda_N(B) = 0$, on aura $f(x) = f(y_0)$ p.p. $x \in A$; si $\lambda_N(B') = 0$, on aura $g(x) = g(y_0)$ p.p. $x \in A$; si $\lambda_N(B) > 0$ et $\lambda_N(B') > 0$, on aura $f(x) = f(y_0)$ p.p. $x \in B'$ et un $y_1 \in B'$ t.q. $f(y_1) = f(y_0)$, $g(y_1) \neq g(y_0)$ avec

$$\left[f(x) - f(y_1)\right]\left[g(x) - g(y_1)\right] = 0 \quad \text{p.p. } x \in A .$$

Ce qui donne $f(x) = f(y_1)$ p.p. $x \in B$ (car si $x \in B$, $g(x) = g(y_0) \neq g(y_1)$) ; puisque $f(y_1) = f(y_0)$, on aura $f(x) = f(y_0)$ p.p. $x \in A$.

8

(i) $A = \left\{(x,y) : 0 < x < 1\,,\ x < y < 1\right\}$;

$$\int_A f = \int_0^1 \mathrm{d}y \int_0^y f(x,y)\,\mathrm{d}x .$$

(ii) $A = \left\{(x,y) : 0 < x < 1\,,\ \sqrt{x - x^2} < y < \sqrt{x}\right\}$

$$= \left\{(x,y) : 0 < x < 1\,,\ \frac{1}{4} - \left(x - \frac{1}{2}\right)^2 < y^2 < x^2\,,\ y > 0\right\}$$

$$\int_A f = \int_{1/2}^1 \mathrm{d}y \int_{y^2}^1 f(x,y)\,\mathrm{d}x$$

$$+ \int_0^{1/2} \mathrm{d}y \left\{ \int_{y^2}^{1/2-\sqrt{1/4-y^2}} f(x,y)\,\mathrm{d}x + \int_{1/2+\sqrt{1/4-y^2}} f(x,y)\,\mathrm{d}x \right\} .$$

(iii) $A = \{(x,y) : 1 < x < 2\,,\ x < y < 2x\}$;

$$\int_A f = \int_1^2 \mathrm{d}y \int_1^y f(x,y)\,\mathrm{d}x + \int_2^4 \mathrm{d}y + \int_{y/2}^2 f(x,y)\,\mathrm{d}x\,.$$

(iv) $A = \{(x,y) : 0 < x < \infty\,,\ x < y < \infty\}$;

$$\int_A f = \int_0^\infty \mathrm{d}y \int_0^x f(x,y)\,\mathrm{d}x\,.$$

10 $1 = \pi/4$.

12 2.

15 (I) 2 ; (ii) $1/3$; (iii) πab.

16 (i) $4\pi abc/3$; (ii) $abc/6$; (iii) $\pi/32$; (iv) $16\,a^3/3$.

CHAPITRE 6

6.1.18

1 (i) 0 ; (ii) $(2a)^{3/2}\pi$; (iii) $6ab$; (iv) $2\pi a^2 b\sqrt{1+b^2}$; (v) $(\sqrt{29}/4)\,\mathrm{e}(\mathrm{e}^4 - 1)$.

2 (i) $1/3$; (ii) $-1/3$; (iii) 0 ; (iv) 0 ; (v) $2\pi^2 - 1/2$; (vi) $26/15$; (vii) 0 ; (viii) 3 ; (ix) $2\pi + 164/3$; (x) $[\mu, \xi, \eta]$.

3 (i) $2\pi a$; (ii) $8a$; (iii) $6a$; (iv) $2\sqrt{2}\pi a$.

5 $8a$.

15 (ii) $x^2y + y^3$; (iv) $(1/2)(x^2+y^2+z^2)$; (v) $a\cdot x$; (vi) $u(x) = \sum_{i=1}^N \varphi_i(x_i)$, φ_i étant une primitive de f_i, $1 \leqslant i \leqslant N$.

16 Masse : $\rho T\sqrt{1+h^2}$; centre de gravité :

$$\left(\frac{1}{T}\sin T,\ \frac{1}{T}(1-\cos T),\ \frac{1}{2}hT\right).$$

17

(i) $\big(1-(-1)^{n+1}\big)/(n+1)$; (ii) $(1/2)\ln 2$; (iii) 0 ; (iv) 2 ; (v) 2π.

6.2.11

1 (i) πab ; (ii) $3a^2\pi$; (iii) $3a^2/2$; (iv) $a^2/2$; (v) $3\pi a^2/2$; (vi) $3\pi a^2/8$.

3 (i) $3\pi/2$; (ii) $-\pi$; (iii) $2(\mathrm{e}^{-\pi} - 1)$; (iv) $1/2$; (v) 56.

5 0 si $(0,0) \notin \Omega$, Ω étant le domaine intérieur défini par γ ; 2π si $(0,0) \in \Omega$.

6 Soit Ω le domaine intérieur déterminé par γ ; si $(1,0) \notin \Omega$ alors l'intégrale vaut 0 ; si $(1,0) \in \Omega$ alors l'intégrale vaut 2π car $\int_\gamma = \int_{\gamma_r}$, γ_r provenant d'un petit cercle de centre $(1,0)$ et de rayon r contenu dans Ω et $\int_{\gamma_r}$ se calcule directement.

9 4/3.

10 0 ;

11 (i) $\pi/4$; (ii) $(1/2)\ln 2$; (iii) 1 ; (iv) 3 ; (v) $\int_0^1 f$, f continue.

6.3.12

2 (i) 14 ; (ii) $(2\sqrt{2})\pi a^2$; (iii) $2\pi(2^{3/2}-1)/3$; (iv)$4a^2$; (v) $2a^2(\pi-2)$; (iv) $8a^2$.

3 (i) $\sqrt{3}/120$; (ii) 9π ; (iii) $2\pi(2^{3/2}-1)/3$; (iv) $\pi R^6/24$; (iv) $\pi/4$.

4 (i) 0 ; (ii) $1/8$; (iii) $\pi R^2 h$; (iv) $\pi R^2(1+R^2/2)$.

5 (i) $1/12$; (ii) $2\pi R^7/105$; (iii) 0 ; (iv) 0.

18 $f^{-1} : S_2 \to S_1$ est donné par

$$f^{-1}(x_1,x_2,x_3) = \left(\frac{x_1}{\sqrt{1+x_3^2}}, \frac{x_2}{\sqrt{1+x_3^2}}, x_3\right) ;$$

$f^{-1} : S_3 \to U$ est donné par $f^{-1}(x_1,x_2,x_3) = (x_1,x_2)$.

6.4.9

1 (i) 0 ; (ii) $12\pi a^5/5$; (iii) 3 ; (iv) $1/8$; (v) $4\pi abc/3$.

6.5.7

1 (i) $-a^2/2$; (ii) $2\sqrt{3}\pi r^2$; (iii) 0 ; (iv) 0 ; (v) $\pi/\sqrt{2}$.

6.6.7

11

(i) Non car $\operatorname{div} F \neq 0$;

(ii) $A_1 = \frac{1}{3}x_3^3 - 2x_1x_2x_3 - \frac{1}{3}x_2^3$, $A_2 = -(x_1^2+x_2)x_3$, $A_3 = 0$ ou

$$A_1 = \frac{1}{4}(x_3^3 - 2x_1x_2x_3 - x_2^3), \quad A_2 = \frac{1}{4}(x_1x_2^2 - x_3x_1^2) - \frac{1}{3}x_2x_3$$

$$A_3 = \frac{1}{4}(x_1^2x_2 - x_1x_3^2 + 2x_1^2x_2) + \frac{1}{3}x_2^3 ;$$

on peut toujours ajouter ∇u à $A = (A_1, A_2, A_3)$, par exemple $u(x) = x_3$ donne $\nabla u(x) = (0, 0, 1)$.

12 $h(x_1, x_2, x_3) = \varphi(x_1, x_2)$, $(x_1, x_2, x_3) \in \mathbb{R}^3$ où $\varphi \in C^1(\mathbb{R}^2)$.

CHAPITRE 7

7.4.8

1

(i) $\partial\Omega = \emptyset$, $R_\Omega = \emptyset$, $S_\Omega = \{0\}$;

(ii) $\partial\Omega = \emptyset$, $R_\Omega = S_\Omega = \{(x_1, \ldots, x_N) : x_N = 0\}$;

(iii) $\partial\Omega = \{x : \|x\| = 1\}$, $R_\Omega = \{x : \|x\| = 1 \text{ ou } 1/2\}$, $S_\Omega = \{x : \|x\| = 1/2\}$;

(iv) $\partial\Omega = \left\{(x,y) : y = f(x),\ 0 < x < \frac{1}{\pi}\right\} \cup \left\{(x,1) : 0 < x < \frac{1}{\pi}\right\}$

$$\cup \{(0,y) : 0 < y < 1\} \cup \left\{\left(\frac{1}{\pi}, y\right) : 0 < y < 1\right\} ;$$

$R_\Omega = \partial\Omega$; $S_\Omega = \left\{(0,0) : \left(\frac{1}{\pi}, 0\right), \left(\frac{1}{\pi}, 1\right), (0,1)\right\}$;

(v) $\partial\Omega = \left\{(x,y) : y = f(x),\ 0 < x < \frac{1}{\pi}\right\} \cup \left\{(x,2) : 0 < x < \frac{1}{\pi}\right\}$

$$\cup \left\{\left(\frac{1}{\pi}, y\right) : 0 < y < 2\right\} \cup \{(0,y) : 1 < y < 2\} ;$$

$R_\Omega = \partial\Omega$; $S_\Omega = \{(0,y) : -1 \leqslant y \leqslant 1\} \cup \left\{\left(\frac{1}{\pi}, 0\right), \left(\frac{1}{\pi}, 2\right), (0,2)\right\}$;

(vi) $\partial\Omega = \left\{(x,y) : y = f(x),\ 0 < x < \frac{1}{\pi}\right\} \setminus \{(x_n, 1) : n = 1, 2, \ldots\}$

$$\cup \left\{\left(\frac{1}{\pi}, y\right) : 0 < y < 1\right\} \cup \bigcup_{n=1}^{\infty} \{(x,1) : x_n < x < x_{n-1}\}$$

où $0 < \ldots < x_n < x_{n-1} < \ldots < x_1 < x_0 = 1/\pi$, $x_n = 2/(4n+1)\pi$, $n \geqslant 1$;
$R_\Omega = \partial\Omega$; $S_\Omega = \{(0,y) : -1 \leqslant y \leqslant 1\} \cup \left\{\left(1/\pi, 0\right), \left(1/\pi, 1\right)\right\}$.

La proposition 7.4.6 est applicable à Ω de (iii).

Bibliographie

[1] A. AMBROSETTI, G. PRODI, *A primer of nonlinear analysis*, Cambridge University Press, Cambridge, 1993.

[2] M. BERGER, *Géométrie* (2 tomes), Editions Nathan, Luçon, France, 1990.

[3] R. BOTT, L. W. TU, *Differential forms in algebraic topology*, Springer-Verlag, Berlin, 1982.

[4] N. BOURBAKI, *Espaces vectoriels topologiques*, Masson, Paris, 1981.

[5] N. BOURBAKI, *Eléments d'histoire des mathématiques*, Masson, Paris, 1984.

[6] TH. BRÖCKER, L. LANDER, *Differentiable germs and catastrophes*, Cambridge University Press, Cambridge, 1975.

[7] R. B. BURCKEL, *An introduction to classical complex analysis*, vol. 1, Birkhäuser, Basel, 1979.

[8] H. CARTAN, *Formes différentielles*, Hermann, Paris, 1967.

[9] G. CHILOV, *Analyse mathématique. Fonctions de plusieurs variables réelles*, Editions Mir, Moscou, 1975.

[10] M. J. CROWE, *A history of vector analysis*, University of Notre Dame Press, Notre Dame, 1967.

[11] G. B. DANTZIG, B. C. EAVES, editors, *Studies in optimization*, Mathematical Association of America, U.S.A., 1974.

[12] R. DAUTRAY, J.-L. LIONS, *Analyse mathématique et calcul numérique pour les sciences et les techniques* (9 tomes), Masson, Paris, 1987.

[13] J. DIEUDONNÉ, *Eléments d'analyse* (9 tomes), Gauthier-Villars, Paris, 1969-1982.

[14] J. DIEUDONNÉ, *Abrégé d'histoire des mathématiques 1700-1900*, Hermann, Paris, 1978.

[15] H. DÖRRIE, *100 great problems of elementary mathematics*, Dover, New York, 1965 ($1^{\text{ème}}$ éd. en allemand, 1932).

[16] J. DOUCHET, B. ZWAHLEN, *Calcul différentiel et intégral* (4 volumes), Presses polytechniques et universitaires romandes, Lausanne, 1983-1989.

[17] L. C. EVANS, R. F. GARIEPY, *Measure theory and fine properties of functions*, C.R.C. Press, Boca Raton, Florida, 1992.

[18] H. FEDERER, *Geometric measure theory*, Springer-Verlag, Berlin, 1969.

[19] W. FELLER, *An introduction to probability theory and its applications*, vol. 1, Wiley, New York, 1957 (2nd ed.).

[20] O. FORSTER, *Analysis 3*, Vieweg, Wiesbaden, 1984.

[21] F. R. GANTMACHER, *The theory of matrices*, vol. 1, Chelsen, New York, 1959.

[22] W. HUREWICZ, H. WALLMAN, *Dimension theory*, Princeton University Press, Princeton, 1941.

[23] K. KNOPP, *Theory and application of infinite series*, Hafner, New York, 1971 (1ère éd. en allemand, 1921).

[24] A. LICHNEROWICZ, *Eléments de calcul tensoriel*, Armand Colin, Paris, 1950.

[25] E. MADELUNG, *Die mathematischen Hilfsmittel des Physikers*, Springer-Verlag, Berlin, 1964.

[26] J. MILNOR, *Topology from the differentiable viewpoint*, University Press Virginia, Charlottesville, 1969 (second printing).

[27] J. R. MUNKRES, *Elementary differential topology* (Annals studies 54), Princeton University Press, Princeton, 1963.

[28] J. R. MUNKRES, *Analysis on manifolds*, Addison-Wesley, Redwood City, California, 1991.

[29] R. NARASIMHAN, *Analysis on real and complex manifolds*, Masson, Paris, 1968.

[30] J. NEVEU, *Bases mathématiques des probabilités*, Masson, Paris, 1970 (2ème éd.).

[31] M. H. A. NEWMAN, *Elements of the topology of plane sets of points*, Dover, New York, 1992 (reprint of 2nd ed., 1951).

[32] A. OSTROWSKI, *Aufgabensammlung zur Infinitesimalrechnung*, vol. III, Birkhaüser Verlag, Basel, 1977.

[33] T. PARTHASARATHY, *On global univalence theorems*, Lectures Notes in Mathematics 977, Springer-Verlag, Berlin, 1983.

[34] G. DE RHAM, *Variétés différentiables*, Hermann, Paris, 1960 (2ème éd.).

[35] A. W. ROBERTS, D. E. VARBERG, *Convex functions*, Academic Press, New York, 1973.

[36] W. RUDIN, *Principles of mathematical analysis*, McGraw-Hill, New York, 1976 (3d ed.).

[37] W. RUDIN, *Real and complex analysis*, McGraw-Hill, New York, 1986 (3d ed.).

[38] L. SCHWARTZ, *Analyse* I-IV (4 tomes), Hermann, Paris, 1991-1993.

[39] M. SPIVAK, *Calculus on manifolds*, Benjamin, New York, 1965.

[40] M. R. SPIEGEL, *Mathematical handbook of formulas and tables*, Schaum's Outline Series, McGraw-Hill, New York, 1968.

[41] W. WALTER, *Analysis* I-II (2 tomes), Springer-Verlag, Berlin, 1985-1990.

[42] D. DE WERRA, *Elements de programmation linéaire avec application aux graphes*, Presses polytechniques et universitaires romandes, Lausanne, 1990.

[43] H. WHITNEY, *Geometric integration theory*, Princeton University Press, Princeton, 1957.

Index

A

adhérer, 24
adhérence, 24
– relativement à, 36
additivité dénombrable, 234
aire, 236, 376
algèbre extérieure, 511
anneau, 236
angle, 6
– polaire, 150
– solide, 400
application, 1, 41
– bilinéaire, 74
– fermée, 41
– ouverte, 41
– propre, 192
approximations successives, 71
atlas, 468
– orienté, 521
argument principal, 150

B

barycentre, 103, 302
– géométrique, 302
base
– canonique, 6
– négative, 9
– négativement orientée, 9
– ordonnée, 6
– positive, 9
– positivement orientée, 9
bord, 400, 411
– régulier, 529
boule
– euclidienne, 21
– fermée, 27
Bourbaki, 330

C

calcul tensoriel, 510
cardinalité infinie, 14
carré ouvert, 21
carte
– locale, 467
– redressante, 206
centre de gravité, 302
cercle, 27
chain rule, 90
champ
– conservatif, 343
– dérivant d'un potentiel, 341
– irrotationnel, 416, 421
– incompressible, 421
– normal unitaire continu, 526
– scalaire, 10, 115, 334
– solénoïdal, 421
– rotationnel, 115
– vectoriel, 10, 115, 334
changement de paramétrage
– direct, 470
– opposé, 470
chemin, 53, 334
– équivalent, 337
– fermé, 335
– homotope, 560
– homotope avec extrémités fixes, 560
– opposé, 337
– polygonal, 54, 335
– rectifiable, 339
– régulier, 335
– régulier par morceaux, 335
– simple, 54, 335
– C^1, 335
– C^1 par morceaux, 335
circulation, 410
clan, 236
codimension, 469
coefficient multinomial, 134
cohomologie de de Rham, 557
colatitude, 169
combinaisons convexes, 67
compacité, 44
complétion, 248
complexe de de Rham, 558

composante, 10
– connexe, 59
composition, 3
condition au bord, 155
condition de
– Dirichlet, 155
– Kuhn-Tucker, 221
– Lorentz, 434
conoïde, 394
constante d'Euler, 328
contraction, 70
contrainte de Kuhn-Tucker, 221
contre-exemple, 275
converger, 28, 30
– absolument, 30
– uniformément, 63
convergence uniforme locale, 112
convolution, 312
coordonnée
– cartésienne, 5
– curviligne, 212, 435
– cylindrique, 169
– locale, 468
– polaire, 150
– sphérique, 169
correspondance, 1
corps
– archimédien, 3
– commutatif, 3
– complet, 3
– ordonné, 3
côtés, 10
courbe, 59, 113, 351
– régulière simple, 204
– orientée négativement, 357
– orientée positivement, 357, 362
C^k-difféomorphisme local, 186
courbe de
– Jordan, 352
– niveau, 114
– Peano, 191
cube, 10
– fermé, 277
– fermé à gauche, 277
– ouvert, 21

D

début, 334
degré, 41
– d'une forme différentielle extérieure, 446
densité
– de mesure, 257
– absolument continue, 257
– gaussienne, 308
– normale, 308
dérivation des fonctions composées, 88
dérivée, 78
– directionnelle, 79
– mixte d'ordre deux, 121
– partielle, 79
– partielle d'ordre k, 122
– partielle de premier ordre, 121
– partielle de second ordre, 121
déterminant, 11
– fonctionnel, 84
détermination principale de l'angle polaire, 150
développement
– limité, 132
– de Taylor, 132
diamètre, 10, 44
difféomorphisme, 167, 173
différentielle, 78
– extérieure, 449
– totale, 91
discriminant, 147
disque
– ouvert, 21, 27
– fermé, 27
distance, 19, 60
– euclidienne, 6, 19
– induite, 19
divergence, 115, 366, 398
domaine, 1, 59
– extérieur, 360, 386
– intérieur, 360, 386
– x-simple, 358
– régulier, 363, 397, 411
– simple, 397, 411
– simplement connexe, 563
– y-simple, 358

E

échange de limite et intégrale, 232
élément
– d'aire, 377
– de volume, 477

ensemble
– au plus dénombrable, 14
– borélien, 233
– connexe, 53
– connexe par arc, 53
– compact, 44
– de cardinalité infinie, 14
– de même cardinalité, 14
– de niveau, 114
– dénombrable, 14
– équipotent, 14
– étoilé, 347
– fermé, 24
– fini, 14
– gaussien, 496
– homéomorphe, 38
– infini, 14
– k-négligeable, 482
– négligeable, 210
– non dénombrable, 14
– précompact, 49
– simplement connexe, 351
– μ-mesurable, 248
énergie
– cinétique, 347
– potentielle, 347
épigraphe, 161
équation
– cartésienne, 374
– implicite, 386
– de Poisson, 425
– des ondes, 434
espace
– borélien, 233
– dual, 503
– hilbertien, 324
– image, 1
– mesurable, 233
– mesuré, 235
– métrique, 19
– métrique complet, 29
– normal, 471
– normé, 18
– tangent, 361, 375, 378, 470
– μ-négligeable, 248
espace de
– Banach, 75
– Hilbert, 75
espérance mathématique, 329
extérieur, 25
extrémité, 334
extremum
– lié, 215
– relatif, 100
– relatif strict, 100

F

facteur intégrant, 456
famille d'éléments dans J indexée pour I, 1
famille cohérente d'orientations, 524
Fermat, 107
fermeture, 24
fibré, 531
fin, 334
flux d'un vecteur à travers une surface, 379, 397
fonction, 1
– à support compact, 313
– bijective, 1
– borélienne, 237
– concave, 70
– continue, 38
– convexe, 67
– croissante, 15
– de carré intégrable, 324
– de classe C^1, 82, 85
– de classe C^2, 121
– de classe C^k, 122
– de classe C^∞, 122
– décroissante, 15
– dérivable, 78
– biharmonique, 125
– différentiable, 78
– harmonique, 125
– indicatrice, 5, 237
– indéfiniment dérivable, 122
– injective, 1
– inverse, 182
– intégrable, 238
– lipschitzienne, 97
– localement constante, 54
– localement intégrable, 313
– localement lipschitzienne, 75, 98
– deux fois continûment dérivable, 121
– k fois continûment dérivable, 122
– mesurable, 236, 237, 250
– négative, 15

– négligeable, 238
– numérique, 5, 238
– partielle, 266
– polynomiales, 41
– positive, 15
– positivement homogène, 92
– rationnelle, 42
– réelle, 5, 238
– réciproque, 182
– simple, 237
– sous-harmonique, 155
– stationnaire, 101
– strictement convexe, 157
– strictement croissante, 15
– strictement décroissante, 15
– strictement négative, 15
– strictement positive, 15
– sur-harmonique, 155
– surjective, 1
– uniformément continue, 38
– vectorielle, 10
– μ-intégrable, 243
– μ-mesurable, 248
fonctionnelle linéaire, 8
fonctionnellement
– dépendant, 211
– indépendant, 211
forme, 8
– alternée, 503
– antisymétrique, 503
– différentielle d'ordre
– différentielle extérieure, 446
– k-linéaire, 503
– linéaire, 503
– multilinéaire, 503
– de Pfaff, 446
– quadratique, 87, 139
formule de
– Cauchy-Frullani, 274
– des accroissements finis, 136
– Euler-Maclaurin, 325
– Green, 404
– Lagrange, 93
– Stirling, 325
– Wallis, 326
frontière, 24
– orientée négativement, 363, 398
– orientée positivement, 363, 398
– non régulière, 488
– régulière, 487
– régulière autour d'un point, 488

G

Galilée, 107
gradient, 79
graphe, 2
groupe
– de cohomologie de de Rham, 558
– fondamental, 565

H

Hardy, 113
homéomorphisme, 38
– local, 186
homotopie, 369, 554
– de lacets, 565
hyperplan d'appui, 160
hypersurface, 469

I

image, 1
– inverse, 2
– réciproque, 453
inégalité de
– Cauchy-Schwarz, 6, 271
– Chebyshev, 249, 271
– Hölder, 323
– Minkowski, 324
– Sard, 285
injection locale autour de, 186
immersion, 204
intérieur, 24
intervalle
– borné, 5
– compact, 5
– fermé, 5
– ouvert, 5, 21
intégrable au sens de Riemann, 229
intégrale
– curviligne, 334
– curviligne du champ vectoriel, 334
– définie, 228
– impropre, 251
– indéfinie, 229
– de Lebesgue, 228
intégration terme à terme, 232

invariance
– de la dimension, 191
– du domaine, 189
invariant, 140
inversion, 170

J

jacobien, 84
juxtaposer, 54
juxtaposition, 343

K

Kettenregel, 90
k-simplexe
– standard, 550
– singulier, 550
k-chaîne singulière, 550

L

lacet, 335
– simple, 335
– homotope, 369
– ponctuel, 563
Landau, 321
laplacien, 167
latitude, 169
lemme de
– Fatou, 240
– Urysohn, 62
lemniscate de Bernoulli, 305
ligne de niveau, 114
limite, 28
– des fonctions, 32
– des fonctions continues, 63
longitude, 169
longueur, 236, 336
loi, 329
– de probabilité, 329
– de la conservation d'énergie totale, 347

M

matrice
– associée, 6
– carrée, 10
– fonctionnelle, 84
– hessienne, 122
– jacobienne, 84
– réelle transposée, 7
matrice de type
– (M, N), 6
– négatif, 139
– positif, 139
– strictement négatif, 139
– strictement positif, 139
– infini, 139
maximum
– absolu, 100
– local, 100
– relatif, 100
– relatif strict, 100
mean value theorem, 93
métrique, 19
mesure, 236
– d-dimensionnelle, 284
– finie, 234
– image, 294
– invariante par translation, 280
– monotone, 234
– naturelle, 480
– positive, 234, 276
– produit, 265
– purement atomique, 278
– superficielle, 376
– μ-intégrable, 243
– σ-finie, 234
mesure de
– comptage, 235
– Lebesgue, 235
– probabilité, 234, 329
– Radon, 276
minimum
– absolu, 100
– local, 100
– relatif, 100
– relatif strict, 100
mineur
– d'ordre k, 12
– d'ordre $(N-1)$, 12
– principal, 146
– principal diagonal, 145
Mittelwertsatz, 93
moment, 302
– angulaire, 302
– d'inertie, 302
monôme, 41
multi-indice, 132
morceau de surface régulière, 373
moyenne, 302
– arithmétique, 222

– géométrique, 222
– harmonique, 223
multiplicateur de Lagrange, 217
multiplication extérieure, 506

N

nombre de
– Bernoulli, 328
– Lebesgue, 74
normale
– dirigée vers l'extérieur, 362, 397, 491
– dirigée vers l'intérieur, 362, 491
– unitaire, 378, 471
normalement convergente
norme
– euclidienne, 6, 18
– équivalente, 23
– maximum, 18

O

opérateur,
– différentiel, 116, 123
– différentiel à coefficients constants, 134
– différentiel linéaire du premier ordre, 124
– différentiel linéaire du second ordre, 124
– hamiltonien, 421
opérateur de
– D'Alembert, 124
– Laplace, 124
– l'équation de la chaleur, 124
– l'équation des ondes, 124
optimisation, 215
origine, 334
orientation, 352, 378, 384
– induite, 531
orthogonal, 6

P

parallélépipède d-dimensionnel, 283
partie
– bornée, 10, 44
– contractile, 556
– convexe, 10,
– discrète, 205
– négative, 242
– ouverte, 21
– positive, 242
– stokienne, 544
paramétrage
– adapté, 529
– approprié, 351
– de même orientation, 521
– d'orientation opposée, 521
– direct, 337
– local, 467
– naturel, 340
– opposé, 337
paramètre, 205
paramétrisation locale, 205
partition, 3
– dénombrable, 3
– de Riemann, 228
– finie, 3
– de l'unité, 460
pavé, 10
– compact, 10
– borné, 10
– borné fermé à gauche, 277
– fermé, 10
– ouvert, 10
permutation, 10
– impaire, 11
– paire, 11
plan tangent, 114
plongement, 204
point, 6
– adhérent, 24
– critique, 101, 209
– critique dégénéré, 148
– critique non dégénéré, 148
– double, 335
– extrémaux
– final, 334
– frontière, 24
– initial, 334
– intérieur, 24
– isolé, 24
– multiple, 335
– régulier, 209, 351
– simple, 335
– singulier, 209
point de
– accumulation, 24, 49
– selle, 143
– stationnarité, 101
point-selle, 13

polynôme de
– Bernoulli, 328
– Taylor d'ordre n, 136
potentiel, 341
– newtonien, 302, 426
– scalaire, 341, 421
– vecteur, 421
précompact
primitive, 228
principe de/du
– maximum, 154, 406
– minimum, 155
– Cavarieri, 265
problème de Dirichlet, 155
produit
– cartésien, 2
– euclidien, 5
– extérieur, 447
– scalaire, 5
– tensoriel, 510
– vectoriel, 13
programmation
– linéaire, 215
– non linéaire, 215
prolongement
– des mesures, 236
– par zéro, 237
propriété de Lindelöf, 288

Q

quaternions, 421

R

rang, 140
– constant, 200
recouvrement, 44
– ouvert, 44
rectangle mesurable, 264
règle d'anticommutation, 446, 447
règle de
– Guldin, 303
– Guldin-Pappus, 393
régularisation, 314
relier, 53
reparamétrage, 337
restriction, 2, 235
rotationnel, 115
ruban de Möbius, 384

S

Schwarz, 387
segment
– fermé, 10
– ouvert, 10
série, 30
– normalement convergente, 64
Shannon, 165
signature, 11, 140
simple connectivité, 192
simple connexité, 369
somme, 30
– partielle, 30
– de Riemann, 228
sous-ensemble
– fermé dans, 36
– fermé relativement à, 36
– non borné, 33
– ouvert dans X, 35
– ouvert relativement à X, 35
– X-fermé ; 36
– X-ouvert, 35
sous-variété
– différentielle de dimension zéro, 205
– différentielle négligeable, 481
– différentielle sans bord, 205
– orientable, 521
sphère, 27
Steiner, 107
submersion, 205
surface, 113
– close, 386
– fermée, 386
– paramétrique, 114
– réglée, 394
– régulière, 204, 383
– orientable, 384
– orientée négativement, 386
– orientée positivement, 386
suite, 4
– convergente, 28
– de Cauchy, 29
– finie, 1
support, 313
– compact, 457

T

tangente, 113
tenseur
– covariant, 178, 510

DATE DUE
DATE DE RETOUR

– covariant a
– covariant d
– contravari
théorème
– fondamen
intégr
théorème d'
– approxima
– Euler, 92
– Helmholtz
– inversion l
théorème de
– Fubini, 23
– fonctions
– Gale et N
– Gauss-Os
– Green, 35
– Hahn et M
– Hadamar
– Lebesgue
– Poincaré,
– Riesz-Ma
– Sard, 210
– Stokes gé
– Stone-We
– Weierstra
théorème d
– converge
– converge
– divergence, 35
– divergence dans $\mathbb{R}^2$, 366
– moyenne, 93
théorème du
– point fixe, 70, 194
– prolongement de Tietze-Urysohn, 65
teorema del valor medio, 93
théorie du degré, 194
topologie, 23
– différentielle, 180
– euclidienne, 23
– induite, 36
Torricelli, 107
trace, 338
trajectoire, 113
– X-ouvert, 35
volume, 236
– N-dimensionnel, 269

W

Weierstrass, 112

X

X-adhérence, 36
x-section, 264

Y

y-section, 264

σ-additivité, 234
μ-presque partout, 248

TRENT UNIVERSITY
0 1164 0441503 0

Cours d'Analyse

1 Analyse vectorielle

Srishti D. Chatterji

L'objectif principal du premier volume de ce Cours d'Analyse en trois volumes est la présentation du théorème de Stokes généralisé pour les sous-variétés différentielles de dimension k dans $\mathbb{R}^N$.
Ce théorème constitue un outil indispensable pour l'analyse dans les variétés et il est une généralisation naturelle des théorèmes dans $\mathbb{R}^2$ et $\mathbb{R}^3$ de Gauss, Green et Stokes; ces derniers étant d'utilisation courante dans plusieurs théories physiques, ils sont présentés d'abord dans le cadre de l'analyse vectorielle dans $\mathbb{R}^2$ et $\mathbb{R}^3$ sous une forme habituellement utilisée par les ingénieurs et les physiciens. Leur généralisation complète dans $\mathbb{R}^N$ exige le recours à la théorie des formes différentielles qui est développée en détail dans cet ouvrage. Toutes les connaissances nécessaires pour comprendre ces développements sont présentées dans les premiers chapitres; elles regroupent les théories de base concernant la topologie et le calcul différentiel dans $\mathbb{R}^N$, les théorèmes concernant les fonctions implicites ainsi que la théorie de l'intégration (de Lebesgue) dans $\mathbb{R}^N$.
Cet ouvrage intéressera tout particulièrement les étudiants en mathématiques et physique du premier cycle universitaire.

Après ses études à l'Université de Lucknow et à l'Indian Statistical Institute de Calcutta, **Srishti D. Chatterji** obtient son doctorat en statistiques mathématiques de la Michigan State University, E. Lansing, USA, en 1960.
Il est depuis 1970 professeur ordinaire de mathématiques à l'Ecole polytechnique fédérale de Lausanne après avoir occupé divers postes d'enseignement et de recherche en Europe, en Amérique du Nord et en Australie.
Ses recherches concernent différents domaines d'analyse liés à la théorie de l'intégration et des probabilités. Il s'intéresse tout particulièrement à l'histoire des mathématiques des 19^e et 20^e siècles et aux fondements de la physique théorique.

ISBN 2-88074-314-1
9 782880 743147

Presses polytechniques et universitaires romandes